Quantitative Health Risk Analysis Methods

Modeling the Human Health Impacts of Antibiotics Used in Food Animals

Recent titles in the INTERNATIONAL SERIES IN OPERATIONS RESEARCH & MANAGEMENT SCIENCE

Frederick S. Hillier, Series Editor, *Stanford University*

Maros/ *COMPUTATIONAL TECHNIQUES OF THE SIMPLEX METHOD*
Harrison, Lee & Neale/ *THE PRACTICE OF SUPPLY CHAIN MANAGEMENT: Where Theory and Application Converge*
Shanthikumar, Yao & Zijm/ *STOCHASTIC MODELING AND OPTIMIZATION OF MANUFACTURING SYSTEMS AND SUPPLY CHAINS*
Nabrzyski, Schopf & Węglarz/ *GRID RESOURCE MANAGEMENT: State of the Art and Future Trends*
Thissen & Herder / *CRITICAL INFRASTRUCTURES: State of the Art in Research and Application*
Carlsson, Fedrizzi, & Fullér/ *FUZZY LOGIC IN MANAGEMENT*
Soyer, Mazzuchi & Singpurwalla/ *MATHEMATICAL RELIABILITY: An Expository Perspective*
Chakravarty & Eliashberg/ *MANAGING BUSINESS INTERFACES: Marketing, Engineering, and Manufacturing Perspectives*
Talluri & van Ryzin/ *THE THEORY AND PRACTICE OF REVENUE MANAGEMENT*
Kavadias & Loch/*PROJECT SELECTION UNDER UNCERTAINTY: Dynamically Allocating Resources to Maximize Value*
Brandeau, Sainfort & Pierskalla/ *OPERATIONS RESEARCH AND HEALTH CARE: A Handbook of Methods and Applications*
Cooper, Seiford & Zhu/ *HANDBOOK OF DATA ENVELOPMENT ANALYSIS: Models and Methods*
Luenberger/ *LINEAR AND NONLINEAR PROGRAMMING, 2nd Ed.*
Sherbrooke/ *OPTIMAL INVENTORY MODELING OF SYSTEMS: Multi-Echelon Techniques, Second Edition*
Chu, Leung, Hui & Cheung/ *4th PARTY CYBER LOGISTICS FOR AIR CARGO*
Simchi-Levi, Wu & Shen/ *HANDBOOK OF QUANTITATIVE SUPPLY CHAIN ANALYSIS: Modeling in the E-Business Era*
Gass & Assad/ *AN ANNOTATED TIMELINE OF OPERATIONS RESEARCH: An Informal History*
Greenberg/ *TUTORIALS ON EMERGING METHODOLOGIES AND APPLICATIONS IN OPERATIONS RESEARCH*
Weber/ *UNCERTAINTY IN THE ELECTRIC POWER INDUSTRY: Methods and Models for Decision Support*
Figueira, Greco & Ehrgott/ *MULTIPLE CRITERIA DECISION ANALYSIS: State of the Art Surveys*
Reveliotis/ *REAL-TIME MANAGEMENT OF RESOURCE ALLOCATIONS SYSTEMS: A Discrete Event Systems Approach*
Kall & Mayer/ *STOCHASTIC LINEAR PROGRAMMING: Models, Theory, and Computation*
Sethi, Yan & Zhang/ *INVENTORY AND SUPPLY CHAIN MANAGEMENT WITH FORECAST UPDATES*

**** A list of the early publications in the series is at the end of the book ****

Quantitative Health Risk Analysis Methods

Modeling the Human Health Impacts of Antibiotics Used in Food Animals

by
Louis Anthony Cox, Jr.
Cox Associates and
University of Colorado

Springer

Louis A. Cox
Cox Associates and University of Colorado
Denver, CO, USA

Library of Congress Control Number: 2005049015

ISBN-13: 978-0-387-25909-3
ISBN-10: 0-387-25909-0
ISBN-13: 978-0-387-26118-8 (e-book)
ISBN-10: 0-387-26118-4 (e-book)

Printed on acid-free paper.

Printed in the United States of America.

9 8 7 6 5 4 3 2 1 SPIN 11349273

springeronline.com

To Cope and Michael –

For making good humor, hard work, and complete commitment to scientific truth central in a political world

Contents

CHAPTER 7
Dynamic Modeling and Uncertainty Analysis

Preface

This book grew out of an effort to salvage a potentially useful idea for greatly simplifying traditional quantitative risk assessments of the human health consequences of using antibiotics in food animals. In 2001, the United States FDA's Center for Veterinary Medicine (CVM) (FDA-CVM, 2001) published a risk assessment model for potential adverse human health consequences of using a certain class of antibiotics, fluoroquinolones, to treat flocks of chickens with fatal respiratory disease caused by infectious bacteria. CVM's concern was that fluoroquinolones are also used in human medicine, raising the possibility that fluoroquinolone-resistant strains of bacteria selected by use of fluoroquinolones in chickens might infect humans and then prove resistant to treatment with human medicines in the same class of antibiotics, such as ciprofloxacin.

As a foundation for its risk assessment model, CVM proposed a dramatically simple approach that skipped many of the steps in traditional risk assessment. The basic idea was to assume that human health risks were directly proportional to some suitably defined exposure metric. In symbols:

$$\text{Risk} = K \times \text{Exposure},$$

where "Exposure" would be defined in terms of a metric such as total production of chicken contaminated with fluoroquinolone-resistant bacteria that might cause human illnesses, and "Risk" would describe the expected number of cases per year of human illness due to fluoroquinolone-resistant bacterial infections caused by chicken and treated with fluoroquinolones.

If it could be made to work correctly, this simple approach would have a lot to recommend it. It obviates the need to develop complex simulation models for exposure or sophisticated dose-response models for human

infection and illnesses caused by foodborne bacteria. It roots the estimate of a single parameter, K, firmly in empirical data on concurrent levels of Exposure and Risk. It provides a simple, easily communicated basis for estimating how a reduction in exposure would affect human health risk.

And yet, there appeared to be potentially fatal conceptual flaws in the approach. First, it implied a positive value of K whenever Risk and Exposure were both positive… and yet, available data showed that handling, cooking, and eating fresh chicken were associated with a *reduced* level of risk for the bacterial illnesses (resistant or susceptible) of concern. How could such a protective effect be reconciled with the model? (Adding an intercept term to the model was one obvious possibility, but even that would not suffice to explain protective effects for low-level exposures and increased risks for relatively rare, high exposures. Perhaps some nonlinearity was needed, but that would threaten the simplicity of the model.) The approach estimated K from historical levels of Risk and Exposure, but did not appear to give valid predictions of how *changes* in Exposure would change *future* levels of Risk… the chief concern for practical risk management. And it relied on exposure metrics, such as prevalence of bacterial contamination in chicken carcasses, that could be estimated from historical data, but that are fundamentally inadequate for predicting risks that depend on the *quantity* of contamination rather than just on its *prevalence*, i.e., its presence or absence. Most of all, the model assumed a direct proportional causal relation between its selected Risk and Exposure metrics without validation, and then focused on quantifying their ratio, K, without validating that the ratio had causal significance or predictive power in real data sets. It thus skipped the traditional (and rather important) role of hazard identification in human health risk assessment.

And yet, the simplicity and intuitive appeal of the approach are strong. Can its technical limitations be remedied and the benefits of simplicity and direct estimation of parameters from data be achieved while allowing valid causal interpretations of the results and useful predictions of human health impacts of changing exposures? This book presents a lengthy, but essentially affirmative, answer. The key steps are to: (a) Allow for other sources of exposure (represented by a "Background" intercept term to be estimated from data); (b) Acknowledge that changes in animal antibiotic use typically affect multiple types of exposures (e.g., reducing resistant bacteria but perhaps increasing susceptible bacteria in processed food commodities); (c) Focus on *causal impacts* on risk arising from *changes* in exposures; and (d) Recognize that the changes in risk predicted by the simple linear model for a given set of changes in exposure can provide valid *bounds* on the true changes that will occur with a more complex but more realistic model. With these enhancements, the mathematical model becomes slightly more

complex. For example, the "Risk = K × Exposure" equation must now be generalized to include additional terms, as in:

> Current Risk = Background + K_1 × (exposure to susceptible bacteria) + K_2 × (exposure to resistant bacteria)

To be interpreted causally, its parameters must be estimated by appropriate statistical methods for causal models, such as structural equations modelling, rather than by straightforward statistical regression. The model's predictions give bounds on true but unknown changes in risk. With these modifications, the goal of obtaining a practical, simple, data-driven approach to estimating probable human health impacts caused by risk management interventions that change exposures to microbial loads can be substantially met.

Many other advances are also possible, such as modeling the systems dynamics of bacteria flows over time between resistant and susceptible types and between ill and well subpopulations of humans and animals. These and other extensions of the basic model are explored in the present book. The methods developed here may also prove useful in other areas of health risk analysis where complex simulation models of risk are difficult or impractical and simpler formulae can produce useful bounds on the human health consequences of alternative risk management interventions.

The public health implications of this new approach to risk assessment are striking. As demonstrated in this book, considering the effects of animal antibiotics on susceptible bacteria as well as on resistant bacteria makes clear that banning current uses of animal antibiotics is *not* necessarily an effective way to preserve the efficacy of human antibiotics or to protect human health, as one might at first suppose. To the contrary, withdrawing current animal antibiotics appears to have the potential to create more ill animals per year, more human illnesses per year, and thus greater use of antibiotics in human medicine, where the opportunity to select resistant strains of bacteria may be far more threatening than in animal uses. Applying a systems dynamics perspective to animal antibiotic use and human health suggests that interventions designed to protect human health by reducing or banning animal antibiotic use may unintentionally increase human health risks. This possibility appears sufficiently strong so that it should be carefully evaluated before policies are made that assume that restricting animal antibiotic use will promote human health.

Since the FDA-CVM, 2001 risk assessment of fluoroquinolones was completed, the world has moved on. Some, including this author, objected to the use of the "Risk = K × Exposure" framework as giving incorrect predictions, being inconsistent with available epidemiological data, and postulating a causal relation where none was apparent in real data. FDA defended its approach, denied the credibility of these technical objections,

and insisted that its risk assessment justified withdrawing approval for fluoroquinolones used in poultry. As this book goes to press, the dispute has proceeded through litigation (with FDA's Administrative Law Judge finding for FDA's CVM in 2004). This litigation clearly demonstrated to both sides and to many stakeholders the need for clearer, sounder, and simpler methods to understand the relevant risks and to improve the basis for risk management decision in this challenging area in the future.

The perspective of this book is that risk is fundamentally quantitative, and when an intervention has some helpful and some harmful effects, it is necessary to use quantitative analysis (at least to place bounds on the sizes of different effects) to compare them and to determine the course of action yielding the largest human health benefits. The major goal of this book is to provide and illustrate methods for quantitative risk assessment and for comparing alternative risk management actions, given realistic limitations on scientific knowledge and available data.

Acknowledgments

First, I thank my colleague Dr. Douglas Popken, who has collaborated with me since 2000 on developing and applying computer simulation models of the human health effects of animal antimicrobials. His tenacity in tracking down real-world data and his creativity and proficiency in simulation modeling led to many of the insights that motivated this book. In addition, Chapters 6 through 8 are based largely on papers that we have written together (Cox and Popken 2004a, 2004b, and 2005).

I thank the journals *Risk Analysis* and *Environment International* for permission to reuse material from published articles. Chapter 1 includes most of Cox *et al.*, 2005, while Chapter 8 is based largely on material from Cox, 2004. The hazard identification framework in Chapter 3 was first sketched in Cox and Ricci, 2005.

The United States Food and Drug Administration's Center for Veterinary Medicine (CVM) first drew my attention to this area and supported part of my review of their risk assessment approach for fluoroquinolones. Their attempts to develop short-cuts and simplifications in traditional risk assessment have served as an inspiration to me to try to figure out how to do the same thing correctly.

Much of the applied research in this book, including development and application of the Rapid Risk Rating Technique (RRRT) approach to quantitative risk assessment, was initially supported by the Animal Health Institute (AHI). Applications to specific animal antibiotics were supported in part by Bayer, Elanco, and Phibro Animal Health. I thank Ken Bafundo, Dennis Copeland, and Tom Shryock for championing the development of better (clearer, more realistic, more valid, more data-driven) approaches to

quantitative risk assessment. I gratefully acknowledge many stimulating conversations with them and with Michael Vaughn and Rich Carnevale on details of animal antibiotics and practical quantitative risk assessment methods. I especially thank Stephen Page for providing detailed, thoughtful comments on Chapters 1-3 that substantially improved the text. I have greatly enjoyed discussing the needs and methods of antimicrobial risk assessment practitioners with Stephen and with Scott Hurd, Gay Miller, Greg Paoli, Paolo Ricci and Randy Singer, as well as occasional correspondence and discussions with Gregg Claycamp. I thank Deborah Doherty and her team at Springer along with Doug Wilcox for an outstanding job in preparing the final text and figures. The methods described here are intended to help animal antibiotic manufacturers, as well as veterinarians and regulators involved in antimicrobial risk assessment, to apply sound, practical, methods of quantitative risk assessment that can be carried out using currently available data while acknowledging (and bounding the effects of) current scientific uncertainties.

Chapter 1

Qualitative and Quantitative Risk Analysis

1. INTRODUCTION: A NEED FOR NEW METHODS

In late 2003, the World Health Organization (WHO), the Food and Agricultural Organization of the United Nations (FAO), and the World Animal Health Organisation (OIE) issued a joint report on risk analysis of human health risks arising from the use of antibiotics in food animals. Written by an expert group that included prominent opponents of animal antibiotic use and regulators from several countries, the report stated that traditional quantitative risk assessment methods are inadequate for antimicrobial resistance risk assessment, due primarily to the uncertainty, complexity and dynamic nature of biological risks and of the evolution and spread of antibiotic resistance among bacteria. It advocated using a more qualitative, expert judgment-based approach instead as a primary basis for risk management decision-making and to categorize animal antibiotics for regulatory action on the basis of human medical need. This approach was to be based largely on precautionary principles and on expert judgments about the importance of drug classes in human medicine, with classes of antibiotics judged "critically important" to be considered for maximal restriction in applications to animals. The report concluded that:

> "Antimicrobial resistance issues crosses (sic) many disciplines, including microbiology, toxicology and pharmacology, and *risk assessment approaches for chemical and microbial contamination are not currently adequate for risk assessment on antimicrobial resistance*. Therefore, when issues pertaining to antimicrobial resistance arise… the questions may need to be referred to a

> WHO/FAO expert body for risk assessment, preferably JEMRA" (WHO, 2003, emphasis added. JEMRA is the Joint Expert Microbial Risk Assessment group that drafted the report.)

It further recommended that:

> "A qualitative approach for risk assessment should be used to make the pre-marketing or postmarketing decision [about approval of animal antibiotics]. Depending on the outcome, the drug sponsor has the option to develop a quantitative risk assessment. ….When dealing with a high level of uncertainty, precaution should be applied in risk management." "The consequences of antimicrobial resistance are particularly severe when pathogens are resistant to antimicrobials critically important in humans. Therefore, the expert workshop recommends that an expert clinical medical group appointed by WHO defines which antimicrobials are considered critically important in humans. … Antimicrobial classes could be classified as critically important when the drug is in a class that is the only available therapy or one of a limited number of drugs available to treat serious human disease or enteric pathogens that cause food borne disease. … Based on the criteria listed previously, a list of critically important classes of antimicrobials should include: the fluoroquinolones and 3rd generation cephalosporins for *Salmonella* spp. and other *Enterobacteriaceae*; the fluoroquinolones and macrolides for *Campylobacter* spp.; and glycopeptides, oxazolidinones and streptogramins for Gram positive bacteria such as *Enterococcus* spp."

These recommendations reflect an increasingly popular movement among many public health experts and health and safety regulators and professionals away from traditional quantitative risk analysis methods and toward more qualitative and judgment-based approaches.

1.1 Challenges and Goals

Qualitative judgment-based approaches to health risk analysis are often perceived as being easier to learn, understand, apply, and explain than quantitative risk analysis methods. They are sometimes also recommended as being more realistic and flexible than quantitative methods in avoiding spurious precision and in dealing with the data gaps and knowledge gaps that inevitably arise for poorly understood and complex health risk phenomena, ranging from human health effects of cigarette smoking (Surgeon General's Report, 2004) to emergence of antibiotic resistance in bacteria infecting human populations to the emergence and spread of BSE among cattle. Indeed, expert judgment-based classification and qualitative characterization of human health risks at first seems appealing for such applications because

of its apparent simplicity, transparency, ease of use, and opportunities for building consensus among groups of experts.

But qualitative judgment-based analyses of public health risks have until now received remarkably little formal evaluation. How well do they work in practice? Do they produce better (e.g., less expensive, more certain, and more effective) risk management decisions than quantitative risk analysis methods? When applied to real data sets, with realistically incomplete, incorrect, and inconsistent information typical of complex health risk applications, do they produce greater or less clarity than quantitative methods? What are the theoretical and practical limitations of qualitative and judgment-based risk assessment? With the exception of related work on the psychology of judgment and decision-making (e.g., Plous 1993), there has been little direct comparison of the performance of quantitative, data-driven risk analysis methods against the performance of more qualitative risk analyses driven by expert judgments. Psychologists have discovered that many lay and expert judgments about causality and risks are often strongly biased by prior beliefs (White *et al.*, 1995) and by cognitive heuristics and biases (Bornstein and Emler, 2001), including the perceived plausibility of envisioned causal mechanisms (Ahn and Bailenson, 1996). These biases can lead to severe under-weighting of empirical evidence and excessive resistance to new evidence and data (Plous, 1993); unawareness of the sources of beliefs and the influence of preconceptions (Fugelsang and Thompson, 2003) and maladaptive causal inferences and decisions. However, in many cases, even relatively simple quantitative methods based on empirical data give more useful and reliable insights and conclusions than expert judgment-based approaches (Grove *et al.*, 2000). The extent to which these insights apply in animal antibiotic risk analysis is an open question.

This chapter reviews past qualitative risk analysis approaches and identifies limitations in the performance of the current generation of qualitative risk rating systems. It appears that all such systems face some important theoretical limitations. Quantitative risk analysis may often be essential for producing information useful for improving risk management decision-making. But, to be most useful, quantitative risk analysis methods must overcome the deficiencies of qualitative approaches and borrow their strengths, especially simplicity and practicality. Bearing in mind current perceptions about the limitations of quantitative risk assessment – especially, that it is inadequate for animal antibiotic risk assessment because of the complexities and scientific uncertainties inherent in the relevant causal pathways and mechanisms (WHO, 2003) – the following chapters therefore seek to develop, clarify, and illustrate principles of quantitative human health risk analysis, using animal antibiotic risks as specific examples.

This book seeks to show that quantitative health risk analysis can be far less limited than its critics sometimes suppose, and that failure to use it

for public health risk management decision-making is likely to produce inferior decisions and fewer public health benefits than could be achieved by using it. The main purposes of this book are:

- To show how to carry out quantitative assessment of human health impacts (both risks and benefits) of alternative risk management interventions using logically and mathematically sound principles of quantitative risk assessment. Although these principles also apply to many other areas of health risk analysis, the examples and applications in this book focus on animal antibiotic use as a particularly important, controversial, and challenging application area where a need for improved risk assessment methods for informing national and international policies has been strongly felt (WHO, 2003).
- To examine and compare the usefulness of qualitative and quantitative risk assessments in informing regulatory decisions about animal antibiotic use;
- To apply quantitative risk assessment methods and to compare results from qualitative and quantitative risk assessment approaches for several antibiotic classes (fluoroquinolones, macrolides, and streptogramins) that are of practical importance in agriculture for preventing and/or treating bacterial diseases in poultry, cattle, swine, or other food animals. All were suggested to be "critically important" for human medicine on qualitative grounds in the WHO report, quoted above.
- To contrast the risk management decisions and probable consequences for public health of using qualitative *vs.* quantitative approaches.

A major theme explored in this book is that public health policy-making is often best served by quantitative risk assessment. Such risk assessment is both practical and essential for informing decisions that will bring about their intended consequences with high probability. Conversely, qualitative and judgmental risk management decisions made without sound quantitative analysis are often too easily swayed by seemingly plausible accounts of how threats to human health *might* work, rather than factual evidence about how they *do* work, leading to perceptions and decisions that give undeserved weight to envisioned hypothetical harms and insufficient weight to real ones. Quantitative risk assessment can help to promote more realistic and effective public health decision-making.

1.2 Overview of Contents

The chapters that follow introduce both qualitative and quantitative risk assessment approaches and illustrate them with simple calculations and examples. This chapter reviews qualitative risk analysis and its theoretical limitations. Chapters 2 through 5 examine the traditional quantitative risk

assessment steps of hazard identification, exposure assessment, exposure-response or dose-response modeling, and risk characterization, including uncertainty and sensitivity analyses. They also explain and illustrate how to do quantitative risk assessment of animal antibiotics using a relatively simple Rapid Risk Rating Technique (RRRT) framework developed for practical applications. Chapters 6 through 8 provide case studies for the streptogramin combination quinupristin-dalfopristin (QD). Fluoroquinolones and macrolides are also addressed briefly in Chapter 8. These three classes of antibiotics have been the focus of many political, regulatory, and activist efforts to remove antibiotics used in human medicine from use in animals.

Throughout the book, a key decision-relevant question addressed is: *What would be the probable human health consequences of withdrawing or restricting current animal antibiotic uses?* Simple analytical models can help to answer this question. Chapters 6-8 also introduce more advanced topics, such as models for long-term emergence of human antibiotic resistance under selection pressure from animal antibiotic uses (Chapter 7).

2. BACKGROUND AND MOTIVATION

2.1 Difficult Tradeoffs

How can antibiotics best be used to promote human health? The answer has been controversial in many countries, in part because of growing fears that using antibiotics now in *any* application, whether treating human patients or promoting growth in food animals, will make antibiotics less effective in *all* applications in the future (APUA, 2002). This creates a perceived conflict between current and future uses, as well as between competing current uses. Various interest groups have created strong national and international pressures to preserve the effectiveness of hard-to-replace antibiotics by reducing uses that they consider non-essential, especially growth promotion in food animals, and eliminating practices such as the widespread supply of prescriptions to consumers who may have viral or other illnesses not likely to benefit from antibiotic prescriptions.

Failing to use antibiotics appropriately now may also increase future health risks, as well as creating preventable present harm to current patients. For example, if failure to treat or prevent an infection in a person or animal allows an infection to spread to others, it may end up causing *more* total antibiotic use than it initially avoids, as well as increasing days of illness for the original victim. Whether it will do so depends on aspects of infectious disease dynamics that are often uncertain and hard to estimate accurately, such as animal-to-person and person-to-person transmission rates of

infectious bacteria; infectious periods and recovery times; and the effectiveness of antibiotics in reducing transmission rates and recovery times and in hastening the emergence and dissemination of resistance.

The following simplified examples illustrate some of the issues and trade-offs involved in assessing how animal antibiotic use (AAU) can affect public health. When an AAU is effective in preventing or reducing animal illnesses, and when these illnesses lead to increased microbial loads in the meat from these animals, then using the animal antibiotic can benefit public health by reducing microbial loads in meat, thus improving the microbial safety of the meat supply. [This can occur, for example, if animals from ill flocks or herds tend to be underweight or in poor condition, and therefore encounter more processing errors and contamination during processing (Russell, 2003; Dawe, 2004).] On the other hand, animal antibiotic use tends to select for antibiotic-resistant bacteria in animals, and such bacteria may pose increased risks to consumers. Which effect on public health dominates requires a quantitative comparison of the reduction in human health risk due to reduced susceptible bacterial loads in meats *vs.* the increase in risk due to increased resistant bacteria. As suggested by the examples, the outcome of the comparison can go either way, depending on quantitative details of how the AAU affects susceptible and resistant illnesses.

Example: Prevention vs. Treatment

Setting: Suppose that introducing a new animal antibiotic would increase the resistance rate to that antibiotic among bacteria in animal food products from 0 to 50% while reducing the total number of human foodborne illnesses per year caused by bacteria on meat servings from ill or underweight animals from 100,000 cases per year to 80,000 cases per year. Each susceptible case of illness causes an average of 6 days of diarrhea and each resistant case causes an average of 8 days of diarrhea (e.g., due to greater virulence of resistant strains).

Problem 1: What is the net human health impact, measured in illness-days per year of diarrhea, of introducing the animal antibiotic?

Solution: Before introducing the animal antibiotic, the average number of illness-days per year in the population from this animal food source is: (100,000 susceptible cases per year) × (6 illness-days per susceptible case) = 600,000 illness-days per year. After introduction, the number of cases falls to 80,000 per year, but each case has a longer expected duration, of (0.5 × 8 + 0.5 × 6) = 7 days instead of 6, due to the postulated 50% antibiotic resistance level among cases. The average number of illness-days per year is therefore: (80,000 cases per year) × (7 illness-days per case) = 560,000 illness-days per year. In this example, the net impact of introducing the

animal antibiotic use is to decrease the average illness-days per year from this food source in the population by 40,000, from 600,000 to 560,000.

Problem 2: In the previous example, suppose that 50% of human illness cases are prescribed a human antibiotic as an empiric treatment, without waiting for a confirmed diagnosis, and that such prescriptions do not affect clinical outcomes for this particular bacterial illness (although they are effective against other types of infectious diarrhea with similar symptoms, motivating the empiric treatment). For simplicity, assume that each prescription directly causes R additional resistant illness case, due to release and spread of the antibiotic (and resistant bacteria selected by it) in sewage or via other pathways. How large would R have to be for the introduction of the animal antibiotic to reduce *resistant* cases per year?

Solution: Let S denote the expected *total* number of resistant cases (direct plus indirect) created per case that is prescribed the antibiotic. Each prescription directly produces R new resistant cases. Each of these has a 50% probability of being prescribed the antibiotic, causing R further cases, and so on. Thus, S satisfies the recursion: $S = R \times (0.5 \times 1 + 0.5 \times S)$. That is, there is a 50% chance of a new resistant case *not* receiving a prescription, so that the branching process terminates with only 1 resistant case (the first) being formed. There is a 50% probability that the case *does* receive a prescription, in which case it generates S expected additional resistant cases. Each prescription produces a total of $0.5R + (0.5R)^2 + (0.5R)^3 + \ldots = 0.5R/(1 - 0.5R)$ expected new resistant cases (for $R < 2$), as may be confirmed by solving the above equation (i.e., $S = 0.5R + 0.5RS$) for S. Before the animal antibiotic is introduced, 50% of 100,000 susceptible cases per year are prescribed the human antibiotic, while afterwards, 50% of 80,000 cases per year are prescribed the antibiotic. The difference is a reduction of $10{,}000 \times [0.5R/(1 - 0.5R)]$ secondary resistant cases prevented per year. Introducing the antibiotic also directly causes $(80{,}000 \times 50\%) = 40{,}000$ additional resistant cases per year. The reduction of $10{,}000 \times [0.5R/(1 - 0.5R)]$ resistant cases exceeds the 40,000 increase in resistant cases if and only if $0.5R > 0.8$. So, introducing the animal antibiotic decreases *resistant* cases in the human population if and only if $R > 1.6$, although it always reduces *total* cases (and illness-days) per year.

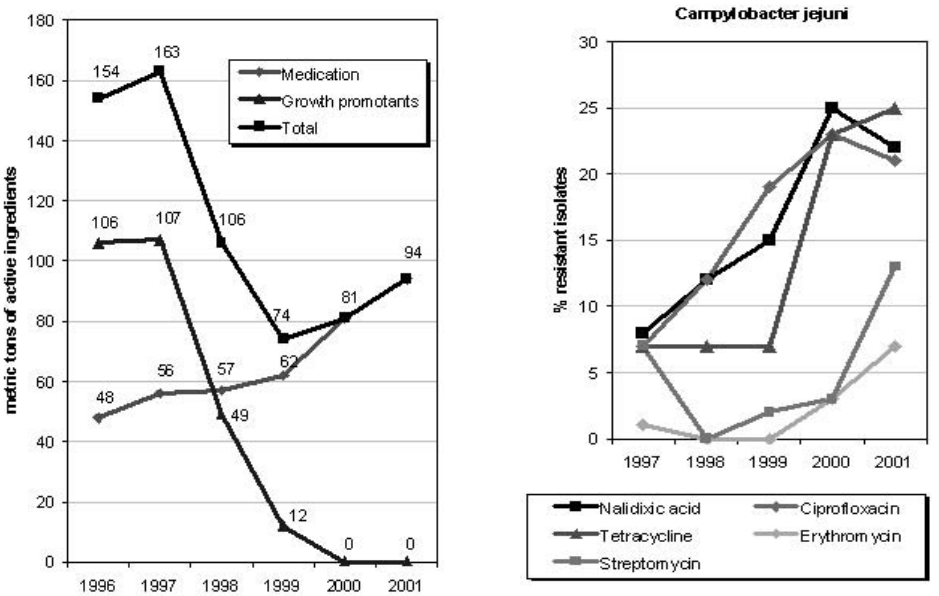
These simple hypothetical calculations show that an exclusive focus on reducing *resistant* cases may sometimes work against the public health goal of reducing the *total* number of foodborne cases and illness-days per year.

2.2 Recent History

Such tradeoffs are not necessarily confined to hypothetical calculations. Figure 1 shows how resistance rates in *Campylobacter* (a common cause of foodborne infectious diarrhea in humans) to various

antibiotics changed following Danish efforts, starting around 1998, to reduce resistance by taxing and effectively banning antibiotic growth promoters (AGPs). Disappointingly, total animal antibiotic use and human antibiotic resistance rates both *increased* dramatically, even if only transiently. Although establishing cause and effect from trend data is challenging, such results, coupled with increases in human foodborne illness rates in many parts of Europe following a European Union ban on antibiotics used as growth promoters (*Eurosurveillance,* 2002), suggest that simply banning animal antibiotic uses is not a panacea for protecting human health in practice. (Bans may not even be a sensible precautionary measure if they *cause* increased human health risks, as discussed further in Chapter 8.) Understanding the quantitative causal relations between animal drug use and human health effects may be necessary to predict the probable human health consequences of bans and other measures and to identify more effective ways to protect and improve human health.

Figure 1: Animal drug use and antimicrobial resistance rates among domestically acquired *C. jejuni* campylobacteriosis cases in Denmark.



Source: Hayes and Jensen, 2003

In any country, the interacting choices and behaviours of many agents – farmers, truck drivers, slaughterers and meat processors, wholesalers, retailers, importers, restaurant kitchen workers and food handlers, inspectors, consumers, institutional workers and inhabitants, patients and doctors and

hospital workers – all affect how many food borne illnesses occur per year and how quickly antibiotic-resistant strains of food-borne bacteria develop and spread. Effective management of these risks requires crossing jurisdictional boundaries and coordinating the interests and interventions of multiple regulatory agencies, decision-makers, and stakeholders. This requirement can be difficult or impossible to meet in practice without huge coordination costs. Coordination and harmonization of food safety standards and animal drug use policies across countries is similarly challenging.

These political and organizational challenges, technical uncertainties, and difficult value trade-offs among competing goals such as food safety, treatment of current human illnesses, prevention of future food-borne illnesses, and ability to treat current and future illnesses effectively, have made it difficult to agree on what policies to follow to manage antibiotic uses and risks of resistance – and perhaps even more difficult to devise policies that will effectively protect and promote human health. Yet, policies now being developed in many countries will affect how antibiotics are used and managed in human and animal medicine for years to come. It is therefore important to determine whether and how to use risk analysis to improve risk management decisions and policies.

Some dramatic actions have already been taken, including phasing out routine uses of animal antibiotic growth promoters in various European countries starting around 1997, and a widely publicized decision by McDonald's corporation in 2003 to buy meat only from producers who do not use them. The human health impacts of such measures are still being studied. So far, it appears that the European interventions to eliminate AGPs were followed by disappointing increases in some human food-borne bacterial illness rates (*Eurosurveillance,* 2002) and antimicrobial resistance rates in clinical isolates from humans with food poisoning (Hayes and Jensen, 2003); by decreases in resistant bacteria in animals and in healthy humans, as hoped (Wegener, 2003); and by transient surges of some antibiotic-susceptible bacterial infections in food animals to unprecedented levels (VLA, 2004). It remains to be seen whether a more science-based strategy, rooted in causal analysis of the relation between actions and their probable human health consequences in the domain of animal antibiotic use, can produce more beneficial results.

There is reason for optimism. From 1996 to the present, even as rates for some food-borne bacterial illness were reaching new heights in Europe, they plummeted in the United States – by over 25% for campylobacteriosis (CDC, 2003). Possible contributors to the decrease may have included: successful implementation of the Hazard Analysis Critical Control Point (HACCP) principles summarized in Box 1 to help identify and control bacterial contamination throughout the food production, processing,

distribution, and commercial preparation chain; increasing public awareness and education; and perhaps continued prudent use of key animal antibiotics by US farmers to help reduce animal illnesses and promote uniform animal weights at slaughter. In Iceland, a concerted effort that included public education, enhanced on-farm biological security measures, and freezing of chicken carcasses, as well as changes in uncontrollable factors (such as the weather) was followed by a more than two-thirds reduction in poultry-borne campylobacteriosis between 1999 and 2000 (Stern and Robach, 2003). Although cause and effect are difficult to unravel in such trend data, it appears plausible that a mix of interventions can reduce human health risks and need for subsequent treatment with antibiotics.

Box 1: Summary of Seven HACCP Principles

- "**Analyze hazards.** Potential hazards associated with a food and measures to control those hazards are identified. The hazard could be biological, such as a microbe; chemical, such as a toxin; or physical, such as ground glass or metal fragments.
- **Identify critical control points.** These are points in a food's production--from its raw state through processing and shipping to consumption by the consumer--at which the potential hazard can be controlled or eliminated. Examples are cooking, cooling, packaging, and metal detection.
- **Establish preventive measures with critical limits for each control point.** For a cooked food, … this might include… minimum cooking temperature and time required to ensure the elimination of any harmful microbes.
- **Establish procedures to monitor the critical control points.** Such procedures might include determining how and by whom cooking time and temperature should be monitored.
- **Establish corrective actions to be taken when monitoring shows that a critical limit has not been met –** for example, reprocessing or disposing of food if the minimum cooking temperature is not met.
- **Establish procedures to verify that the system is working properly** – for example, testing time-and-temperature recording devices to verify that a cooking unit is working properly.
- **Establish effective recordkeeping to document the HACCP system.** This would include records of hazards and their control methods, the monitoring of safety requirements and action taken to correct potential problems. Each of these principles must be backed by sound scientific knowledge: for example, published microbiological studies on time and temperature factors for controlling foodborne pathogens.

Source: USDA/FDA, 2004; http://www.cfsan.fda.gov/~lrd/bghaccp.html

2.3 The Role of Risk Analysis

Risk analysis provides a framework for systematically identifying the factors and processes that affect human health risks from antibiotic use in food animals and for partitioning them into manageable components that can be modeled and quantified using available data. The component modules or sub-models can be combined to give an overall quantitative risk assessment model for predicting and/or explaining the probable human health impacts of different risk management interventions. Validated risk assessment models can also be used to help identify risk management policies and interventions that correspond to desirable predicted risk profiles – those with relatively high human health benefits and low human health losses.

While such risk analysis information cannot by itself solve the political and coordination problems of antibiotic risk management, it does provide the technical information needed to identify which interventions are most likely to protect and promote human health, and which are least likely to do so.

3. QUALITATIVE RISK ANALYSIS

Qualitative risk analysis increasingly provides the foundation for practical risk rating systems and regulatory guidance and requirements documents used in international trade, food safety, and health risk assessment work, including regulatory analyses of the human health risks associated with food animal antibiotic uses (AAUs). These systems assign ratings such as “High”, “Medium”, or “Low” to dimensions of exposure and potential harm and then combine these component ratings to determine an overall rating of risk, to be used as an input to decision-making. Qualitative risk rating systems simplify risk assessments by reducing the required inputs and calculations to a manageable set of judgments, while making the rating logic transparent and easy to apply. They usually require only a few qualitative judgments as inputs, together with supporting reasoning and documentation, and usually produce simple categorizations of risk as outputs that can be communicated easily to policy makers and stakeholders.

3.1 Some Examples of Qualitative Risk Rating Frameworks

Motivated in part by concerns that quantitative risk assessment of human health risks from animal antibiotic use (AAU) might prove to be overly burdensome to implement, produce insufficiently credible or excessively assumption-dependent conclusions, and/or require data that are not readily available or else have to make assumptions of doubtful validity to

bridge important data gaps, several regulatory risk analysis groups worldwide have proposed qualitative rating approaches designed to avoid these pitfalls. For example, a three-component risk rating with components of "Hazard", "Exposure" and "Impact" has been developed in Australia to assist in assessing and characterizing the risks associated with resistant bacteria from animal antibiotic use. Risk is profiled with the help of the following matrix:

Qualitative Risk Assessment Framework from Australia (Each factor is rated N = negligible, L = low, M = medium, or H = high)

Factor	Definition
Hazard = source of risk	Antibiotic resistant microorganisms or their resistance plasmids (that have the potential to transfer to humans) within an animal species, arising from the use of an antibiotic in an animal species
Exposure	Amount and frequency of exposure of susceptible humans to antibiotic-resistant microorganisms (or their plasmids) from animal sources
Impact	The evaluation of infections (caused by antibiotic-resistant pathogens of animal origin) in susceptible humans. Considers: a) Perceived or known clinical importance of the class of antibiotics to humans; b) Dose-response assessment of relationship between frequency and magnitude of exposure of humans (dose) to antibiotic-resistant food-borne microorganisms and severity and/or frequency of the impact (response); including an estimate of the critical threshold of exposure required to cause infection in susceptible humans. c) Antibiotic-resistant disease severity, morbidity, mortality. d) Expected numbers of infections and deaths. e) The impact on human health and quality of life including the range of the susceptible humans expected to be affected. Probability of antibiotic-resistant infection development in susceptible humans (N = negligible, L = low, M = medium, H = high)

Source: Adapted from Australia National Registration Authority Veterinary Requirements Series, Part 10 http://www.apvma.gov.au/guidelines/vetguideline10.pdf.

This framework is used as part of "the general requirements for submitting antibiotic resistance data for the registration of veterinary chemical products that contain antibiotics as active constituents", specifically "any proposed use in Australia of a product containing a new antibiotic" or "any proposed extension of use in Australia of a registered product containing an existing approved antibiotic where the NRA [National Registration Authority] considers that there is likely to be a significant increase in the volume of usage or that there may be an increased risk to public health as a result of the use of that antibiotic" (Australian Pesticides

and Veterinary Medicines Authority, Part 10 of Veterinary Requirements Series, www.apvma.gov.au/guidelines/vetguideline10.pdf). Assessments in this framework may include separate narrative risk summaries for different bacterial species and discussions of uncertainty in the supporting data used in the risk assessment and of possible benefits of use of antibiotic in Australian animal health (so that these impacts may be considered in parallel with risks of adverse impacts). "Risk" is characterized as "Probability of disease due to infection in susceptible humans after exposure of humans to antibiotic-resistant microorganisms (and genetic material) of animal origin **and** the severity of the impact of exposure on susceptible humans". This conceptualization of risk, although referring to probability, omits details typically included in quantitative risk assessments, such as a specified time interval or denominator (e.g., per year or per capita-year) for the "probability of disease"; a clear distinction between "after" and "caused by"; and specification of how the number of susceptible humans (or the causes of susceptibility) are to be included in assessing risk. The qualitative "impact" category contains items (e.g., dose-response relation and clinical importance of human antibiotics) that might be redundant for quantitative risk assessment, once the expected number and severity of additional morbidities and mortalities caused by a change in AAU are known. However, the framework contains many ideas that are useful in any risk assessment, including consideration of expected illnesses and deaths; distinctions among illnesses of different severities; and identification of (perhaps multiple) susceptible subpopulations and multiple bacterial species if required to adequately characterize risk.

In Ontario, Canada, risk analysts at the Ontario Ministry of Agriculture and Food (OMAF) (McNab and Alvas, 2003) have commented as follows on the relation between qualitative and quantitative risk assessment: "Quantitative risk assessments are preferred. Unfortunately, detailed quantitative data are frequently not available. In such cases, the OMAF framework strongly encourages the organization of qualitative assessments in a format that is aligned with quantitative risk assessment." They propose a qualitative rating system for risk analysis of various threats, including bacteria, using H = high, M = medium, L = low, N = negligible for risk, its components, and its impacts. The process is described as follows:

> "When reliable quantitative data is available, assessors use quantitative multiplicative mathematical models to estimate risk. Often, the desired quantitative data are not available. In such cases a more qualitative approach is used. In either case, quantitative and qualitative assessments are summarized using a rating system to help categorize risks. The final rating assigned to a given hazard/commodity situation is derived from six sub-

ratings, each rated as negligible, low, medium or high. The first three sub-ratings are concerned with the probability of a human health impact being realized. This is influenced by several factors including the exposure characteristics of the situation. The final three sub-ratings are concerned with the impact of the disease, which is influenced by several factors including dose-response characteristics. This scoring system is used to help categorize risks in terms of their general importance. It is not used to rank individual risks in numerical sequence, but does attempt to place them in broad categories of negligible, low, medium or high risk."

The proposed framework has the attractive feature of comparing probabilities of consequences with and without different risk management interventions. It uses multiplication to aggregate the components of the risk rating when adequate data are available, appropriately implying that a microbial hazard that creates negligible human exposure or for which exposure has negligible adverse human health impact can have a risk rating of negligible even if other factors are large.

An important aspect of this system is that it considers the *changes* in estimated probabilities of risk components if different risk actions are taken. This concept – using risk rating systems to link proposed risk management actions to their probable consequences, defined as changes in the probabilities (or of statistical frequencies in affected populations) of the outcomes of interest – can be applied to many settings. In particular, it suggests that the human health risk of a proposed change in AAU, such as introduction of a new product or withdrawal of an existing one, should be assessed by considering how human health impacts are likely to change if the proposed action is taken. This emphasis on the human health consequences of risk management decisions is consonant with many recommendations that risk analyses should be decision-focused and provide information useful for assessing risk management decision options.

The Brenner Scheme for order-of-magnitude risk rating

The UK's Brenner scheme for genetically modified organisms (GMOs) (http://www.hse.gov.uk/biosafety/gmo/acgm/acgmcomp/2a1.pdf) assigns to each of the following three factors an order-of-magnitude weight (such as 10^{-3}, 10^{-6}, or 10^{-9}):

- ACCESS = probability that the GMO, or DNA contained within, it will be able to enter the human body and survive.
- EXPRESSION = measure of the anticipated or known level of expression of the inserted DNA

- DAMAGE = measure of the likelihood of harm being caused to a person by exposure to a GMO, *independently* of access and expression

Multiplying these three factors yields an overall risk factor, which is used to look up a provisionally recommended containment level for a GMO experiment. Although this approach is considered to be only one component to help inform risk management, not a comprehensive risk assessment approach by itself, the following ideas may be applicable outside the GMO context to AAU risk analysis: using order-of-magnitude numerical estimates of risk factors or risk components (see Darwiche and Goldsmitz, 1994 for a formalization of order-of-magnitude probabilistic reasoning via the "kappa calculus"); multiplying the results to get a rough quantitative estimate of overall risk; and mapping this rough quantitative risk estimate into a provisional risk management decision category.

US FDA-CVM's Guidance Document # 152 and its limitations

In 2003, the US FDA (FDA, 2003) provided a guidance document on qualitative risk assessment to the animal antibiotic industry, Guidance Document #152. This serves as the default approach for organizing and presenting risk information in seeking approval for new animal antibiotics in the US. The FDA Center for Veterinary Medicine (CVM) considers qualitative risk assessments provided in accord with Guidance #152 in making decisions about whether and under what conditions to approve new drug applications (NDAs) and product line extensions for animal antibiotics.

The guidance document conceptualizes risk as a *probability* of human illness (over an unspecified time frame and population) caused by certain antimicrobial-resistant bacteria of interest and treated by resisted antimicrobials. Specifically, it defines risk as *"The probability that human food-borne illness is caused by an antimicrobial resistant bacteria (sic), is attributable to an animal-derived food commodity, and is treated with the human antimicrobial drug of interest."*

Example: Some Limitations of Qualitative Risk Definitions

Question: What are the limitations of the preceding definition of risk?

Answer: The preceding definition of risk has the following limitations:

- It does not specify the *frequency* of the risk event. "The probability" referred to in the definition lacks a specified denominator (e.g., probability per capita-year, or per century for the whole exposed population, etc.) Yet, an intervention that reduces the *frequency* of occurrence from one occurrence per person-year to one

occurrence per million person-years should be considered to have reduced risk, even if the *probability* that some cases occur each year remains constant at 1.

- It ignores human health consequences and the *severity* component of risk. By this definition, a bacterium that has a 99% probability of causing 1000 deaths per day could have a lower "risk" (i.e., lower probability) than one that has 100% probability of causing one excess day of mild diarrhea in the population per century. "Risk" refers to both probability (or frequency) *and* magnitude (or severity) of harm. Any definition of risk should incorporate both components.
- It omits any reference to human health harm. For example, if treatment with the antimicrobial drug of interest were always completely effective (e.g., because the normal therapeutic dosage is sufficient to kill even resistant bacteria), then risk should be less than if all treatments are always ineffective for resistant bacteria. Similarly a very mild illness poses less risk than one that is very severe or fatal, other aspects (e.g., frequency of occurrence) being the same.
- The phrase "attributable to an animal-derived food commodity" is ambiguous. For example, suppose that the resistant bacteria originate in sewage from hospitals, but reach farm animals in drinking water and are found on meat products from the animals. Should such bacteria be considered "attributable to an animal-derived food commodity"? In practice, "attributable to" is often interpreted to mean "is statistically associated with". But there may be no causal relation between a risk and the source(s) to which it is attributed. For example, consider the causal graph: $Z \leftarrow Y \leftarrow X \rightarrow W \rightarrow V \rightarrow \text{Risk}$, where an arrow from one variable to another indicates that the first causes the second. If the empirical relations among these variables are that $Z = Y$, $Y = X$, $\text{Risk} = V$, $V = W$, and $W = X$, then the "attributable risk" as defined in epidemiology will be *identical* for all five variables (and each will be equal to 100%, if all variables are 0-1 binary variables), even though only X is exogenously determined and only V is a direct cause of Risk. Thus, the risk *attributable* to a variable (such as 100% for Z) does not necessarily reflect the risk *caused* by it (which is zero for Z, in this example).
- The definition does not consider *causality*. If the presence of the resistant bacteria in the specified food commodity is not affected by the animal antibiotic use (AAU) of concern, for example, then it should not be considered part of the risk from that AAU. For example, suppose that banning the AAU would cause *more* resistant bacteria to be ingested per capita-year by susceptible humans (e.g., because the AAU is not the source of resistant bacteria, which instead reach animals via drinking water, flies, wild birds, or other pathways; and ceasing the AAU would amplify microbial loads, including the resistant portion, reaching consumers.) Then the mere presence of such bacteria in an animal-derived food commodity, which is all that the definition requires, should not be considered part of the risk attributed to the AAU.

Like the other qualitative systems, Guidance #152 rates risks based on ratings of their components – in this case, "Release", "Exposure", and "Consequence" components. Following the related conceptual framework proposed at a World Health Organization workshop WHO (2003), the "Consequence" component refers not to actual human health consequences, but to qualitatively judged "importance" of drugs in human medicine. Unlike the Australian and Ontario systems, the US Guidance #152 approach does not include a separate zero or "negligible" (N) qualitative value. Risk is always rated "High" if the consequence component is rated "Critically important", even if there is negligible or zero release potential, exposure, and adverse human health consequence from exposure. Thus, the risk rating is dominated by the Consequence rating.

The criteria for being "critically important" proposed by WHO 2003 include having only a limited number of available alternatives and being used to treat human foodborne illnesses. On these grounds, WHO suggests that important classes of animal antibiotics including streptogramins (which CVM labels as "highly important"), fluoroquinolones, and macrolides should all be categorized as "critically important", and therefore "High" risk. A limitation of this approach is that it neglects information on *how often* these drugs are used. As it turns out, other drugs are often preferred to and used instead of these for the specific bacterial food-borne illnesses of concern. [Thus, linezolid (Zyvox™) is increasingly used instead of the streptogramin combination Synercid™ for vancomycin-resistant *E. faecium* infections. Antimicrobial therapy for the few exceptionally severe cases of campylobacteriosis that most warrant antibiotic treatment (e.g., in bacteremic and immunocompromised patients) may begin with gentamicin, imipenem, third-generation cephalosporins, or chloramphenicol until susceptibility test results are available, rather than with macrolides or fluoroquinolones (Ang and Nachman, 2003). However, Benson *et al.* (2005) suggest that, "As with non-HIV-infected patients, the optimal treatment of campylobacteriosis among persons with HIV-1 infection is poorly defined. Among patients with mild disease, certain clinicians might opt to withhold therapy unless symptoms persist for more than several days. Increasing resistance to fluoroquinolones makes the choice of therapy especially problematic. For mild-to-moderate disease, initiating therapy with a fluoroquinolone (ciprofloxacin) or a macrolide (azithromycin), pending susceptibility test results, and treating for 7 days is a reasonable approach".] In general, the human health consequences of resistance cannot be estimated from the above criteria for judged human medical importance of each antibacterial drug, as these criteria do not include how often the drugs are prescribed (e.g., always *vs.* never) to cases with resistant bacteria.

The risk rating directly determines suggested risk management action categories (e.g., strictly limited use, intermediate restriction, or least restriction). Antibiotics classified as "High" risk are identified as typically subjected to the most restrictive use conditions (e.g., "strictly limited" use), without considering whether such limitations benefit or harm human health or whether different risk management options would better protect human health. In general, as emphasized in the Ontario approach, rational risk management decision-making should consider the *changes* in frequencies and severities of human health effects caused by proposed risk management interventions. Since Guidance #152 does not consider such changes, it may be less useful for decision-making than other approaches. However, this reflects limitations of the particular approach, rather than an intrinsic limitation of all qualitative approaches.

3.2 Lessons from Previous Approaches

Comparing the risk rating systems above suggests several worthwhile ideas and components to include in any (qualitative or quantitative) risk rating system for animal antimicrobials. A list of potentially valuable concepts to include in future risk rating systems follows.

1. ***Frequency* of adverse human health impacts.** Frequency of harm should be part of the definition of risk in any system, quantitative or qualitative. As discussed further in Chapter 2, the conceptual units of frequency are *expected number of illnesses per year* (in an identified exposed population), for population risks; and *expected number of illnesses per capita-year* for individual risks.

Technical Note: Population heterogeneity. Frequencies should ideally be estimated for relatively homogeneous subpopulations, i.e., subpopulations whose members have approximately equal risks; otherwise inter-individual heterogeneity in risks must be addressed. Statistical techniques such as classification tree analysis (Zhang and Singer, 1999; Lemon *et al.*, 2003) and finite mixture distribution modeling can help to identify homogeneous subpopulations and to estimate frequencies for them from case-control, cohort, and longitudinal survey data. Cox, 2001, especially Chapter 3, provides a survey of relevant statistical techniques for risk analysts.

2. ***Severity* of adverse human health impacts** from different risk management decisions (e.g., proposed changes in AAU). As suggested by the Impact portion of Australia's qualitative risk assessment template, the severity of human health impacts from preventable illnesses should be a key component of the risk assessment. The conceptual units for severity are *expected adverse impacts per illness* (e.g., mortalities,

morbidities, illness-days, life-years lost, etc.), perhaps with morbidities further broken down by severity class (e.g., mild, moderate, severe) and with mortalities further classified by age group or number of life-years lost. *Quality-adjusted life-years (QALYs)* lost may also be used if the required assumptions (Hazen, 2003; Miyamoto, 1999) are accepted and if it is desired to aggregate diverse health impact metrics into a single summary measure. (As with frequency, severity of health impacts should be assessed for multiple subpopulations, e.g., based on age, immune status, etc., if impact severity distributions differ significantly among them. Classification tree analysis and other modern statistical methods can help to identify relevant subpopulations from data.)

The motivation for considering severity of human health impacts in rating risks is illustrated by the following example. Suppose that "probability of human illness caused by a specified resistant bacteria, attributable to a specified animal-derived food commodity, and treated with the human antimicrobial drug of interest" = 1, but that treatment with the human antimicrobial drug of interest is completely effective clinically (i.e., resistance makes no difference to clinical outcome). This situation should presumably be rated as less severe (lower risk) than one in which the corresponding probability is less than 1 but the impact is treatment failure and death due to resistance. To assure that the second situation is rated as more severe, human health impacts must be considered.

3. ***Causality* for adverse human health impacts created by proposed *changes* in animal antibiotic use (AAU).** As suggested by the Ontario approach, it is useful to be able to assess the *change* in expected adverse human health consequences per year or per capita-year as a result of proposed risk management interventions.
4. ***Uncertainty*** about the changes in frequency and severity of adverse human health effects caused by a proposed risk management intervention. For example, what overall rating should be assigned to a situation that has a 50% chance of an "L" risk rating, a 30% chance of an "M" rating and a 20% chance of an "H" rating, depending on how scientific uncertainties are resolved? In the Ontario system, uncertainty is summarized along with risk characterization information before a final overall risk rating is applied. In the Brenner system, uncertainty about the component ratings is indicated by order-of-magnitude estimates and these uncertain estimates are then used to identify risk management responses.
5. ***Cumulative risk assessment,*** i.e., total risk summed over the multiple pathways by which effects of risk management decisions (e.g., changes in animal antibiotic use) accumulate to cause resulting changes in adverse human health effects. These pathways may include multiple

bacterial species and/or multiple drugs with co-resistance or cross-resistance; multiple food products; and perhaps multiple human subpopulations affected. They must typically also include susceptible as well as resistant strains of bacteria if both are affected by proposed changes, so that total human health impact can be considered.

6. ***Potential risk reduction benefits*** to humans and animals. To inform rational risk management, both reductions and increases in risk from proposed interventions must be assessed. Animal health benefits can also be listed separately in the overall assessment of impacts of proposed risk management interventions, as described in the Australian system.
7. ***Necessary and sufficient information.*** The systems considered list many potentially relevant and informative data elements to be considered in the rating process. Exactly how these data elements should be assembled to build up a coherent account of the overall human health risk caused by a propose change in AAU is less clearly specified. It is therefore possible that several overlapping or partly redundant pieces of information that address essentially the same bottom-line concern (e.g., exposure, response probability, etc.) might be considered while leaving unaddressed other key information (e.g., on the human health impacts specifically caused by resistance-related treatment failures) needed for decision-makers to understand how changes in AAU will affect human health risks.
8. ***Multiplicative aggregation.*** As stated by McNab and Alves, 2003 and many other sources, *multiplication* is appropriate for aggregating suitably defined quantitative components to form an overall risk estimate. As an example, suppose that the following quantities can be estimated, perhaps to the nearest order of magnitude:
   - *Exposure factor = ΔExposure* = (change in contaminated servings ingested per year) caused by the change in animal antibiotic use, ΔAAU. (If the dose-response relation increases from near zero below a "minimal infectious dose" threshold to near 1 above it, then a "contaminated" serving is one that carries at least the threshold number of bacteria, i.e., enough to increase the probability of illness in a susceptible individual. If illness probability is approximately proportional to number of bacteria ingested, with no threshold, then the exposure factor is the number of bacteria ingested per serving.)
   - *Dose-Response factor = (ΔIllnesses/ΔExposure*) = (expected number of additional illnesses) per (contaminated serving ingested) (or per bacterium ingested, for linear no-threshold dose-response functions.)
   - *Consequence factor = (ΔHuman health impacts/ΔIllnesses)* = expected number of adverse health consequences per illness case resulting from ingestion of a contaminated serving. If multiple

impacts are considered, then separate consequence factors can be estimated for the different types of impacts (e.g., illness-days by severity category, mortalities, QALYs lost, etc.)

Then for a given change in animal antibiotic use on the farm, the corresponding human health risk would have an estimated value determined by the product:

Risk = Exposure factor × *Dose-Response factor* × *Consequence factor*

where the variables on the right-hand side are the factors just described. The conceptual units of risk are change in adverse human health consequences per year (or per capita-year, for individual risks) in the exposed population from the proposed change in animal drug use.

This product is appropriate for a single *combination* of the exposure, dose-response, and consequence factors, e.g., for a specific animal drug, bacterium, strain (susceptible or resistant), food commodity, exposed susceptible subpopulation, and adverse effect category. To estimate total risks, it is necessary to sum the risks over all combinations in the intended scope of the risk assessment. Thus, multiplicative aggregation of component estimates is natural for each combination, while additive aggregation is natural across combinations.

Technical Note: Combinations may be thought of as cells of a large contingency table (or as leaf nodes in a classification tree) of factor combinations determining expected illnesses per capita-year for exposed individuals. Given the number of individuals in each cell (its "*size*") and the estimated expected illnesses per capita-year for individuals in that cell (its "*risk*" rate), the expected total illnesses per year in the population is the sum over all cells of the *size* × *risk* product. The probability distribution of total illnesses will be approximately Poisson, and hence determined by the expected number of illnesses. The sum-of-products framework is useful for uncertainty analysis, as products of uncertain factors tend to be approximately log-normal, sums of uncertain products are approximately normally distributed, and sums of products may be insensitive to specific numbers (Henrion *et al.*, 1996).

Example: Calculating Quantitative Impacts of a Proposed Ban

Setting: Suppose that the human health impacts of banning a particular animal antibiotic are estimated to be as follows:

- *Exposure factor:* Average number of infectious bacteria ingested per serving of the food commodity increases by a factor of 1.09 due to increased average bacteria contamination per meal following the ban. The proportion of resistant bacteria declines from 15% to 0%.

- *Dose-response factor:* Expected number of illnesses per year caused by ingesting contaminated servings of the animal food commodity is directly proportional to the amount of infectious bacteria ingested per serving (i.e., the dose-response relation is a linear no-threshold relation.) Since exposure increases by 1.09, so do expected illnesses per year. (The linear no-threshold dose-response relation justifies use of average microbial load per serving as the exposure variable, rather than fraction of servings with at least a certain level of contamination, for example.)
- *Consequence factor:* Each resistant illness causes an average of 8 days of diarrhea. Each susceptible illness causes an average of 6 days of diarrhea.

Problem: What is the human health impact of the ban, under these assumptions? (Compare the change in steady-state illness-days per year in the population before and after the ban, without attempting to model the transient adjustment process or indirect effects due to use of human antibiotics to treat illness-days.)

Solution: In this example, there are only two groups of bacteria, susceptible and resistant, to be summed over. It is convenient to handle the summation by simply taking weighted averages of appropriate terms. Before the ban, the average value of the Consequence factor is: 0.85 × 6 days + 0.15 × 8 days = 6.3 days. After the ban, this decreases to 1.00 × 6 + 0 × 8 days = 6 days, assuming that the ban is completely effective in eliminating all resistant cases from this food source. Thus, the illness-days-per case decreases by a factor of (6/6.3) = 0.9524. On the other hand, the ban increases the expected total number of illnesses per year by 9%, due to increased exposures to microbial loads in food. In this model, illnesses are proportional to microbial loads, with the proportionality factor, which is the Dose-Response factor, remaining the same before and after the ban. Population risk is proportional to *Exposure Factor* × *Dose-Response Factor* × *Consequence Factor*; therefore, the net result is an increase in risk from this food source of: 1.09 × 0.9524 = 1.038, i.e., an approximately 4% increase in illness-days per year (and per capita-year) in the population, provided that other factors (e.g., consumption and cooking habits) remain unchanged.

Mathematical justifications for multiplication as the way to combine component values can be found in multiattribute value and utility theory (e.g., Hazen, 2003 for QALYs); in dimensional analysis (e.g., Buckingham's Pi Theorem) and fundamental measurement theory (Luce and Suppes, 2001); or in applied probability laws for decomposing joint probabilities and expected values into products of marginal and conditional probabilities and expected values. To apply multiplicative aggregation to qualitative ratings, it is necessary to have a zero or negligible rating (such that products of rated factors are negligible whenever at least one of the factors is). In fact, as developed next, even with such zero ratings, there is in general no

mathematically sound way to combine ordered categorical ratings to mimic multiplication. This will provide an incentive to consider stronger rating scales, including those used in quantitative risk analysis approaches, while still allowing for realistic uncertainties in the input component values.

4. A MATHEMATICAL THEORY OF QUALITATIVE RISK RATING

Risk rating systems such as those above can be formally modeled by mathematical functions $z = f(x_1, x_2, \ldots, x_n)$, or $z = f(\mathbf{x})$, where $\mathbf{x}$ is a vector of input attribute levels to which a risk level or rating is to be assigned and $z = f(\mathbf{x})$ is the risk rating assigned to $\mathbf{x}$. In the qualitative rating systems reviewed, a detailed description $\mathbf{x}$ is first mapped to a coarse description, $\mathbf{y} = g(\mathbf{x})$, where $\mathbf{y}$ is a small (e.g., three-component) set of labels for qualitative risk components such as release, exposure, and consequence. Each component of $\mathbf{y}$ is a qualitative label (e.g., H, M, or L). The component rating vector $\mathbf{y}$ is then assigned a final risk rating label z (e.g., H, M, or L) by some other function, h, which is often represented in practice as a look-up table. The whole rating process may thus be summarized by the following diagram: $\mathbf{x} \xrightarrow{g} \mathbf{y} \xrightarrow{h} z$. In a quantitative risk assessment, the quantitative risk associated with situation $\mathbf{x}$, denoted by $r(\mathbf{x})$, is a numerical function of its quantitative attributes, i.e., the components of $\mathbf{x}$. These attributes (e.g., exposure, dose-response potency, and consequence attributes) are typically oriented so that more is worse. As just discussed, the quantitative risk associated with $\mathbf{x}$ is often a product of such numerical factors, or a sum of such products over multiple hazards, exposure pathways, and exposed individuals.

Most qualitative risk rating systems use *direct qualitative ratings* of quantitative factors, defined as follows. The domain of each quantitative factor is partitioned into consecutive contiguous intervals, each of which is assigned a qualitative label from an ordered set of labels. Each qualitative label corresponds to an interval of values for the quantitative factor. A label's interval lies to the right of all the intervals for lower-ranked labels and to the left of the intervals for all higher-ranked labels. Generally, a direct qualitative rating system may use different sets of qualitative labels for different components of $\mathbf{x}$, and the cardinalities of these label sets may differ. We assume that all components of $\mathbf{x}$ have been oriented so that higher values are worse (i.e., risk is a non-decreasing function of its components, for both qualitative and quantitative ratings). This assumption will be called *monotonicity*. L and H will denote the least and greatest of the ordered qualitative labels, respectively, in the range of f.

A basic consistency requirement for qualitative and quantitative risk assessments is *soundness*, which states that higher quantitative risks should receive higher qualitative risk labels, or, at least, should not receive lower ones. The Appendix proves the following mathematical result:

THEOREM 1: No direct qualitative rating system satisfying monotonicity is sound for arbitrary quantitative risk functions, or even for those functions of greatest practical interest, such as the product function, $r_p(\mathbf{x}) = x_1x_2...x_n$ = *product of components of* $\mathbf{x}$.

In other words, given any direct qualitative rating system $f(\mathbf{x})$ and a quantitative risk function such as $r_p(\mathbf{x}) = x_1x_2\ldots x_n$, it is always possible to choose two inputs, say $\mathbf{x}$ and $\mathbf{w}$, such that $\mathbf{x}$ is assigned a higher *qualitative* risk rating than $\mathbf{w}$, even though $\mathbf{x}$ has a lower *quantitative* risk than $\mathbf{w}$. The following simple numerical example illustrates the problem.

Example: Qualitative vs. quantitative risk reversals

Suppose that each of three components of $\mathbf{x}$ has discrete possible values of 0, 1, 2, or 3, and that these quantitative values are mapped to corresponding qualitative labels as follows: g(0) = L, g(1) = g(2) = M, g(3) = H. Then, if quantitative risk is a product of the component values, the quantitative vector $\mathbf{x}$ = (1, 1, 3) has a qualitative rating of $g(\mathbf{x})$ = (M, M, H) and a quantitative risk of $r_p(\mathbf{x}) = 1*1*3 = 3$. The vector $\mathbf{w}$ = (2, 2, 2) has a qualitative rating of $g(\mathbf{w})$ = (M, M, M) and a quantitative rating of $r_p(\mathbf{w}) = 2*2*2 = 8$. By monotonicity of qualitative ratings, $\mathbf{x}$ must be assigned a qualitative risk rating $f(\mathbf{x})$ that is at least as high as $f(\mathbf{w})$ (since $g(\mathbf{x}) \geq g(\mathbf{w})$ component-by-component) even though its quantitative risk is less than half as great. Thus, $f(\mathbf{x}) \geq f(\mathbf{w})$, even though $r(\mathbf{x}) < r(\mathbf{w})$, i.e., *the qualitative and quantitative ratings are inconsistent.* The theorem shows that such contradictions always exist, i.e., they cannot be eliminated by using more rating levels or by more careful coding of quantitative values as qualitative labels. Thus, for example, if the qualitative coding is changed to: g(0) = g(1) = L, g(2) = M, g(3) = H., then by monotonicity of qualitative ratings, f(1, 1, 3) = h(L, L, H) ≤ h(L, M, H) = f(0, 2, 3), i.e., f(1, 1, 3) ≤ f(0, 2, 3) even though the quantitative risks are $r_p(1, 1, 3) = 1*1*3 = 3$ for the first and $r_p(0, 2, 3) = 0*2*3 = 0$ for the second. Thus, again, $f(\mathbf{x}) \geq f(\mathbf{w})$, even though $r(\mathbf{x}) < r(\mathbf{w})$, i.e., the qualitative and quantitative ratings are inconsistent (where $\mathbf{x}$ = (1, 1, 3) and $\mathbf{w}$ = (0, 2, 3)). In general, *no* direct qualitative rating system satisfying monotonicity can represent the product risk function.

Some qualitative risk rating systems now in general use (FDA, 2003) add additional layers of qualitative rating, e.g., by basing the exposure component of the rating process on qualitative ratings of subsidiary factors, as shown in Table 2. Other components likewise are calculated from

subsidiary components by qualitative rating tables, often expressed in terms of corresponding contiguous intervals of quantitative values, such as "Low (< 5%), Medium (5-25%), High (> 25%)".

Table 2: Qualitative Rating of Probability of Human Exposure

	Amount of food commodity being consumed		
Amount of food commodity contamination	**High**	**Medium**	**Low**
High	H	H	M
Medium	H	M	L
Low	M	L	L

Source: FDA, 2003, Table 5, page 19

Such indirect or *hierarchical* rating systems may be diagrammed as follows: $\mathbf{x} \xrightarrow{q} \mathbf{v} \xrightarrow{g} \mathbf{y} \xrightarrow{h} z$. Here, a detailed quantitative attribute vector **x** is first mapped to an array **v** of qualitative attribute values (such as High, Medium, or Low for the attributes "Amount of food commodity contamination" and "Amount of food commodity being consumed" in Table 2). This mapping is denoted by q in the diagram. Then, **v** is mapped to a higher-level set of qualitative attribute values, such as H, M, or L for "Probability of Human Exposure" in Table 2. This mapping is denoted by g. Finally, possibly after several such layers of qualitative mapping and aggregation, the top-level qualitative attribute value vector **y** (such as H, M, and L values for release, exposure, and consequence qualitative attributes) is mapped to a qualitative risk label (such as H, M, or L) via another look-up table, denoted by h (e.g., Table 6, p. 22 of FDA, 2003).

Successive layers of qualitative coding can introduce loss of information and inconsistency in the interpretation of labels. The following example illustrates that *numerical probabilities cannot be assigned consistent qualitative labels* by partitioning the unit interval [0, 1] into sub-intervals (each corresponding to a qualitative label such as H, M, or L) that allow the qualitative labels to be consistent both with the underlying quantitative probabilities and with commonly used rules for combining qualitative labels.

Example: Qualitative Aggregation Rules Inconsistent with Probabilities

Consider a two-layer risk-rating hierarchy with two first-level variables (e.g., risk = x_1x_2, where x_1 = probability of exposure and x_2 = conditional probability of an adverse consequence given exposure), each derived from two subsidiary (lower-level) input variables (e.g., x_{11} = probability of purchasing a contaminated serving and x_{12} = probability of not cooking it adequately, for exposure; and x_{21} = conditional probability of infection given ingestion of a contaminated serving and

x_{22} = conditional probability of illness given infection, for consequence). Suppose that aggregation of qualitative ratings satisfies the following *unanimity* condition: if all of the subsidiary variables from which a higher-level variable is derived have the same qualitative value (e.g., H, M, or L), then the aggregate rating of the higher-level variable has the same qualitative value as its inputs. (Thus, as in Table 2, the diagonal elements have the same qualitative labels as the corresponding row and column variables.) Finally, suppose that some specific desired interpretation of qualitative labels is intended for the input probabilities and the output probability ("risk" = x_1x_2) in terms of ranges of corresponding quantitative values, such as "Low = probability < 5%; Medium = probability between 5% and 25%; High = probability > 25%". Since all of the probabilities in question, x_1, x_2, x_{11}, x_{12}, x_{21}, and x_{22} refer to probabilities of adverse events, we assume for simplicity that the same quantitative-to-qualitative coding is to be used for each of them (although the argument can be generalized). In general, in a hierarchical risk rating system satisfying unanimity, *no such consistent quantitative interpretation of qualitative labels is possible.* The reason is that products of quantitative probabilities that belong to the lowest extreme of the "Medium" range, for example, should not receive the same qualitative value as the product of probabilities at the upper extreme of the "Medium" range, and similarly for other ranges. For example, set all four inputs (x_{11}, x_{12}, x_{21}, x_{22}) equal to 0.26 in the above example. Then the quantitative risk for the product function $r_p(\mathbf{x})$ is $r_p(0.26, 0.26, 0.26, 0.26) = (0.26)^4 = 0.0046$, corresponding to a qualitative probability label of L. But unanimity implies that f(0.26, 0.26, 0.26 0.26) = h(H, H, H, H) = H. Thus, the qualitative and quantitative risk ratings give opposite results.

Hierarchical aggregation of qualitative labels even in this small example results in upward-biased qualitative risk ratings that do not allow a qualitative rating of H to discriminate between quantitative risks of 0.0046 and 1. If each of the four inputs x_{11}, x_{12}, x_{21}, x_{22} is independently sampled from the unit interval U[0, 1], then simulation shows that a qualitative risk rating of H is highly likely to be mistaken (probability of almost 95%) when compared to its intended quantitative definition. By increasing the number of inputs in this example, the error probability can be made arbitrarily close to 1 (and the sizes of quantitative risks that are labeled "H" by any rating system incorporating unanimity can be made arbitrarily close to 0).

Example: Qualitative vs. Quantitative Risk Ratings in Practice

The potential discrepancies between qualitative and quantitative risk ratings are of more than purely theoretical interest. As previously noted, qualitative reasoning and risk ratings of fluoroquinolones, streptogramins, and macrolides have suggested that all three are "critically important" in human medicine and so should be rated as "High" risks typically recommended for maximally restricted use (WHO, 2003, p. 17). But quantitative assessments using $r_p(\mathbf{x})$ models with uncertain input values represented by upper bounds (for human health risks) and lower bounds (for human health benefits), as discussed in Chapter 8, suggest that banning uses of these

drugs in chickens, for example, is expected to cause at least hundreds of excess illness-days for each illness-day prevented for enrofloxacin; thousands of excess illness-days per illness-day prevented for macrolides; and tens of thousands of excess cases of campylobacteriosis per *E. faecium* infection treatment failure prevented for virginiamycin, due to increased animal and human bacterial illnesses caused by terminating current uses. Thus, qualitative and quantitative approaches can lead to very different risk management recommendations in practice, as well as to contrasting results in theory.

How much of this discrepancy between qualitative and quantitative results is intrinsic to the use of qualitative rating methods, as opposed to particular implementations? For example, would including a "Negligible" category, adding more rating levels, or changing the number of attributes used necessarily reduce the discrepancy between qualitative and quantitative risk rating results? As shown below, the general answer is that no change in how individual attributes are rated qualitatively can guarantee that a qualitative risk rating system will give accurate or useful results.

4.1 Results for Uncertain Inputs

An often-perceived potential advantage of qualitative over quantitative risk rating systems is that their inputs (i.e., qualitative ratings of inputs using labels such as H, M, and L) better reflect the rough, imprecise, but useful knowledge available in practice than do overly-precise numerical inputs. This section examines how well qualitative rating systems perform in the presence of *uncertain* inputs. While many qualitative systems do not specify how to rate uncertain inputs (e.g. what label to assign to an input that is judged to have a value of "H" with probability 25% and a value of "L" otherwise), mathematical constraints can be used to bound the performance of *any* system that rates each input separately and then combines these ratings to determine an overall rating of risk.

Qualitative rating systems in widespread use, including all the ones reviewed earlier in this chapter, require rating each component of risk separately. None considers the *joint distribution* of component values in assigning values to each component. However, in general, the discrepancy between qualitative and quantitative results depends on the joint distribution of attribute values. As illustrated in the following examples, neglecting statistical dependencies among inputs can leave a rating system unable to distinguish between quantitative risks that differ by arbitrarily large amounts, e.g., between probabilities of 0 and 1 for an adverse event.

Example: Marginal Distributions of Inputs Do Not Determine Risks

Suppose that risk depends on only two inputs, x_1 and x_2, and that the quantitative relation is: Risk = $r_p(\mathbf{x}) = x_1x_2$. For example, the variables could be interpreted as: risk = illness probability, x_1 = Pr(exposure > 0) and x_2 = Pr(illnesss | exposure > 0). Let the two input values be uncertain, with x_1 and x_2 each being equally likely to be 0 or 1 for each individual and their values across different individuals being statistically independent. (Thus, only about half the population is susceptible.) *Any* rating system that assigns an overall risk rating based only on this information (i.e., on the marginal distributions of the inputs) omits information essential for correct risk assessment. Thus, if $x_2 = x_1$ (susceptibility and exposure are perfectly positively correlated), each individual has a risk of 0.5 = Pr(x_1 = 1); whereas if $x_2 = 1 - x_1$ (susceptibility and exposure are perfectly negatively correlated), then each individual has a risk of 0. A rating system that ignores such correlation information must assign the same risk rating to these very different situations; the assigned ratings, therefore, may not be very informative about the true risk when the components of risk are correlated.

Analogous examples with different risk functions (e.g., $r(\mathbf{x}) = \max(x_1, x_2)$ and $r(\mathbf{x}) = \min(x_1, x_2)$) show that qualitative risk ratings that depend only on the marginal distributions of the inputs must assign the same ratings to joint distributions of $\mathbf{x}$ giving quantitative risks as low as 0 or as high as 1, depending on the correlations among components (because their marginal distributions are the same).

A similar lack of discriminatory power can be demonstrated even if no quantitative risk function is considered. Suppose that qualitative labels are assigned to risk factors and combined according to the pattern in Table 2 (i.e., component ratings of (H, H), (H, M) and (M, H) are assigned risk ratings of H; (L, L), (L, M) and (M, L) are assigned risk ratings of L; and (L, H), (M, M) and (H, L) are assigned risk ratings of M.) Then distributing all probability density uniformly over the three cells (L, L), (M, H) and (H, M) gives the same uniform marginal distributions for the qualitative ratings of each component as distributing it uniformly over the three cells (H, H), (M, L), and (L, M). In other words, the first joint distribution, which gives a 2/3 probability of an H risk rating and a 1/3 probability of an M risk rating, would have to be assigned the same qualitative risk label as the second joint distribution, which gives a 1/3 probability of an H risk rating and a 2/3 probability of an L risk rating, by any procedure that rates each component based only on its own marginal distribution of probabilities for qualitative values. Yet, the first distribution clearly dominates the second.

4.2 Other Possibilities for Qualitative Risk Rating

The preceding analysis examined qualitative risk analysis approaches that attempt to assign ordered categorical values (labels) to risks and to their

components. Other qualitative and semi-quantitative approaches might use different evaluation scales, such as a rank-ordering of risky prospects from most- to least-risky, or assignment of intervals on a "risk scale" to uncertain prospects (Bilgic, 1997; Neapolitan, 1991; Davidson and Ryks, 2003). Mathematical analysis can help to identify the limitations of what any risk rating system can achieve.

For example, suppose that a rating system is to be used to compare two alternatives, A and B, to determine which should be ranked higher in competing for scarce risk-management resources or regulatory concern. If the overall rating of risk is to be based on component ratings developed for several risk components or factors, as in all of the above examples, then how should the overall risk rating of alternatives A and B depend on the component ratings? Some apparently reasonable properties might include the following.

Box 2: Possible Desiderata For Qualitative Rating Systems

1. Which of alternatives A and B is rated higher in the overall risk rating should depend *only* on their component ratings. Thus, the components used to rate risk should be sufficient to do the job: together, they should determine whether A is assigned a higher, equal, or lower rating than B.
2. Which of A and B is rated higher on overall risk should be able to depend on *each* of their component ratings. Specifically, if A and B are identical in all respects except that A rates higher or worse than B on one factor (e.g., exposure), then B should not be rated higher than A in the overall risk rating. This property should hold for all the risk components: none of them is irrelevant.
3. If A rates higher (or worse) than B on *every* component rating, then B should be rated no higher (or worse) than A in the overall risk rating. For example if A involves greater exposure, more illnesses, and more severe consequences than B, then A should receive a risk rating at least as high as B's.
4. Risk ratings of A and B should be based *only on their own data,* i.e., whether A is rated higher or worse than B should not depend on what other alternatives (other than A and B) are also being rated, if any.
5. If one or more component ratings are zero (e.g., for exposure potential or for human health impact potential of exposure), then the overall risk rating should be zero (or "Negligible" in systems with that category).
6. If the rating for a component is uncertain (e.g., if it has a 0.2 probability of being "L", 0.5 probability of being "M", and 0.3 probability of being "H"), then the single "equivalent" rating assigned to it (i.e., H, M, or L after considering its uncertainty) should not depend on the ratings assigned to the other components.

Although such logical relations among the component ratings and the overall risk rating may be desirable, they can impose strong constraints on possible rating systems. For example, if quantitative ratings are used, then conditions such as 5 and 6 imply that the aggregation formula used to combine component ratings into an overall risk rating must be *multiplicative,* i.e., the overall risk rating is proportional to a product of its component ratings (see Miyamoto, 1999 for details). Such multiplicative aggregation of quantitative ratings satisfies properties 1-6. On the other hand, if only ordinal rankings are used for the components, then Arrow's impossibility theorem implies that there is no qualitative ranking system that can assign coherent overall risk rankings (meaning complete, transitive rank-orderings with ties allowed), based on arbitrary component rank-orderings, in such a way that apparently reasonable principles such as 1-4 are satisfied. Similar limitations may hold for aggregating fuzzy ratings of linguistic labels or scales (e.g., H, M, L, and N), depending on how they are formalized (Bilgic, 1997). In other words, qualitative component ratings may not contain enough information to be coherently aggregated into an overall qualitative risk rating that is related to them in desirable ways.

Another possible concern is that a risk rating system with only a few possible outcome categories may not produce enough information to make a good decision if it is *not complex enough* to adequately reflect the input information. The minimum amount of complexity in the input-output mapping required for a classification system (including a risk rating system) to make few errors can be rigorously analyzed via techniques from information theory and computational learning theory (see e.g., Goldman, 1991, Chapter 7 and Burges, 1998). A key insight from such analysis is that a classification system that lacks enough complexity to discriminate well among essentially different situations may lead to poor decisions, i.e., ratings with high error rates and high expected losses from decision errors.

In summary, additional mathematical analysis approaches, perhaps including axiomatic and complexity-theoretic methods, may provide additional insights into the possible properties and limitations of qualitative risk analysis systems. However, the present generation of qualitative approaches is based primarily on assignment and aggregation of ordered labels, as analyzed above. The theoretical limitations of such systems suggest that it would be premature to abandon quantitative risk analysis if quantitative (especially, multiplicative and sum-of-products) risk assessment models can be made to work well in practice by giving easily calculated, robust, correct results. That is the central task of the following chapters.

5. WHAT SHOULD BE DONE INSTEAD?

The preceding examples have explored some basic limitations of qualitative risk analysis systems, emphasizing that they can perform poorly even in simple situations, such as when quantitative risks are well described by a product of factors. To improve the usefulness of results in such cases, *simple quantitative risk models should be used instead of qualitative risk models.* For example, in general, any joint probability density of random variables $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ (which might represent risk and its components) can be factored into the product form:

$$\Pr(\mathbf{x}) = \Pr(x_1) \times \Pr(x_2 \mid x_1) \times \ldots. \times \Pr(x_n \mid x_1, x_2, \ldots, x_{n-1}).$$

This is of the product form $r_p(\mathbf{x})$ if we re-define the components of $\mathbf{x}$ to be the respective conditional probabilities. To deal with uncertainties, risk analysts often use conservative (upper-bound) estimates of the components of such products, recognizing that doing so gives estimates of $\Pr(\mathbf{x})$ that may be too high but that are unlikely to be too low. Rough but useful upper-bound estimates can often be obtained from currently available data, at least in applications to animal antimicrobial risk assessment (see Chapters 6-7). The product model $r_p(\mathbf{x})$ then gives an upper-bound estimate on the risk from continuing a current practice such as an animal antibiotic use. If desired, analogous calculations can be used to obtain a lower-bound estimate for the risk of *not* continuing the current practice. If the upper-bound estimate of the risk of continuing is much smaller than the lower-bound estimate of the risk of ceasing the current practice, then continuing is a risk-minimizing strategy. Similarly, if an upper-bound estimate of the risk of ceasing is much smaller than the lower-bound estimate of the risk of continuing, then ceasing minimizes risk. If neither condition holds, more information is needed to make a confident choice that is robust to the relevant uncertainties.

In summary, simple quantitative models such as product-form models (or, more generally, comparisons of sums and differences of products) with data-driven upper-bound and/or lower-bound estimates of components of the products will often be more accurate and more useful than qualitative risk ratings, while requiring no more information than would be needed to assess, justify and interpret qualitative ratings. Instead of estimating and comparing qualitative ratings such as H, M, and L, it may be more practical and more meaningful to estimate, combine, and compare interval-valued estimates for risk and its components (Neapolitan, 1991; Davidson and Ryks, 2003).

6. DISCUSSION AND CONCLUSIONS

This chapter has presented some initial results toward a formal mathematical analysis of the performance of qualitative risk rating systems. The main conclusion is that such systems have potentially important limitations, both in theory and in practice. They can create ranking reversal errors (assigning larger qualitative risk ratings to quantitatively smaller risks) for some pairs of inputs and produce qualitative risk ratings that have no clear quantitative meanings and that provide little or no information about true (quantitative) risks.

How well qualitative rating systems work in practice depends on the joint distributions of the components being rated. For example, if cases can be clearly separated into three clusters, with the risks (and risk components) in cluster A all being larger than those in B which are all larger than those in C, then qualitative rating using H, M, and L can discriminate perfectly among these three clusters. Similarly, a qualitative rating system in which a rating of "Negligible" for any component (such as Exposure or Consequence) reliably implies that the total quantitative risk must be small enough to lie below an action threshold, while ratings of "High" for all components reliably imply that action is needed, might serve as a useful and economical screening test that makes further quantitative analysis unnecessary in some cases (WHO, 2003). However, qualitative rating may perform extremely poorly for problems that do not naturally cluster in a way that justifies reliable qualitative ratings. This underscores the importance of carefully evaluating the performance of proposed risk assessment approaches (using mathematical analysis or simulation or empirical test sets or a combination of approaches) before encouraging their widespread use.

This chapter has also introduced several needs for improved human health risk analysis methods identified in the application area of animal antibiotic use. These needs focus on practicality of input requirements, simplicity and clarity of calculations, and ease of explanation of outputs. They have motivated several qualitative approaches to health risk analysis. But the theoretical criticisms of qualitative approaches just summarized suggest that, despite their desirable motivations and simplicity, such approaches may often produce misleading results and/or fail to provide essential information needed to improve the choice among risk management interventions with uncertain health consequences.

These results suggest that it is important to continue to develop and apply practical quantitative risk assessment methods for broad classes of situations in which qualitative methods are not necessarily reliable. The following chapters therefore seek methods of risk assessment that retain the advantages of simplicity while producing more informative and useful – usually quantitative – answers.

APPENDIX: PROOF OF THEOREM 1

For simplicity, assume that the quantitative attributes of risk, i.e., the component dimensions of **x**, are scaled so that each component has a minimum value of 0 and a maximum value of 1; thus, the set of possible **x** values, denoted by X, is the unit cube. This is formalized by the following condition:

Monotonicity: The qualitative risk rating function f(**x**) is a non-decreasing function of its arguments, with f(**0**) = L and f(**1**) = H, where L and H denote the least and greatest of the ordered qualitative labels, respectively, in the range of f. [Here, **0** denotes the vector in X consisting of all zeroes and **1** is the vector in X consisting of all ones. f is the qualitative risk rating functions mapping qualitative component values (i.e., ordered categorical labels) to risk value.] Similarly, the component-rating and risk-rating functions g and h are non-decreasing function of their arguments.

A qualitative rating function f will be called a *sound* representation of a quantitative risk function r if it satisfies the following consistency condition expressing compatibility between f and r:

Soundness: f is a *sound qualitative representation* of the quantitative function r if and only if, for any two quantitative vectors **w** and **x** in X, f(**w**) ≥ f(**x**) if r(**w**) ≥ r(**x**).

Thus, to be sound, f must not assign **x** a higher qualitative risk label than **w** if **x** has a smaller quantitative risk than **w**. (Here, ≥ denotes numerical ordering for r values and ordering of the qualitative labels for f values.)

Consider the product function, $r_p(\mathbf{x}) = x_1x_2\ldots x_n$ = product of components of **x**. Theorem 1 states that no direct qualitative rating system satisfying monotonicity can give a sound qualitative representation of this simple quantitative function.

THEOREM 1: No sound, monotonic, direct qualitative rating representation exists for the product function $r_p(\mathbf{x})$ on the unit cube $X = [0, 1]^n$, for $n > 1$.

Proof: We first give the proof for the unit square, n = 2, with X = [0, 1] × [0, 1]; see Figure A1. Suppose that there were a sound, monotonic, direct qualitative rating representation. Then X would be partitioned into a grid of rectangular cells by the partitions of x_1 and x_2 into contiguous intervals that are assigned the same qualitative labels (i.e., by lines of the form $x_1 = m_i$, $x_2 = m_j$, where m_i is the boundary point between contiguous intervals i and i + 1 for x_1 and m_j is the boundary point between contiguous intervals j and j + 1 for x_1 and indices i and j range over the sets of ordered qualitative labels for x_1 and x_2, respectively.) Let (x, y) > (0, 0) be the lower left corner point and let (u,v) > (x, y) be the upper right corner point of some cell that is labeled H, Figure A1. (The upper right-most cell, corresponding to f(**1**)= H, is one such cell.) Then f(x + ε, y + ε) = H for any feasible ε > 0, such as point ♦ in Figure A1, while for all sufficiently small ε > 0, $r_p(x + \varepsilon, y + \varepsilon) \le r_p(x - \varepsilon, v - \varepsilon)$ (since xv > xy), as at point * in Figure A1. Thus, for f to be a sound representation of r_p, it must be the case that f(x – ε, v – ε) ≥ f(x + ε, y + ε) = H, and so f(x – ε, v– ε)

= H. Thus, if any cell receives a qualitative risk rating of H, then soundness and monotonicity imply that any cell immediately to the left of it must also receive a rating of H. By a symmetric argument, the cell directly below it (if any) must also receive a rating of H. Iterating, all cells must receive a rating of H. But this contradicts the requirement that f(**0**) = L. Thus, assuming that a representation satisfying the conditions of the Theorem exists leads to a contradiction, showing that no such representation exists. For arbitrary n > 1, the proof is similar: starting from the "top right corner" cell, corresponding to f(**1**) = H, for f to be a sound representation of r_p, the cells adjacent to it must be labeled H. This argument is continued until the contradiction f(**0**) = H is obtained. <>

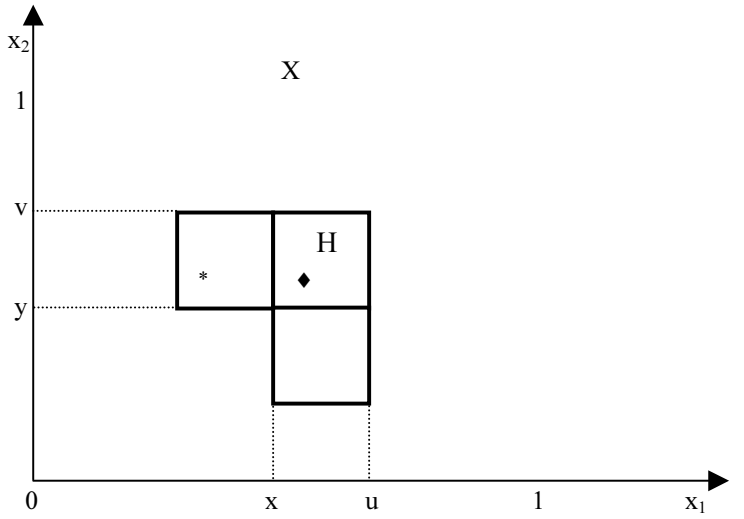


Figure A1: Geometry of Theorem 1

The above proof depended only on the existence of a "top-most" cell (weakly outranking all others on all components) corresponding to f(**1**) = H and on the existence of a "bottom-most" cell (weakly outranked by all others on all components) corresponding to f(**0**) = L, and on comparing the lower left corner of each cell mapping onto H to the upper right corner of the cell to its left (or below it). Hence, it holds for any compact (closed, bounded) set, X, provided that the definition of "monotonic" is extended to imply that top-most and bottom-most cells receive different qualitative labels. Moreover, it can be extended from the product function to any continuous function that increases in all of its arguments. The equal-value contours of any such function slope downward (from upper left to lower right). In the neighborhood of any point where such a contour intersects the midpoint of a cell edge between two adjacent cells, there will be points above the contour curve (i.e., numerically greater risk) and other points below it (i.e., numerically smaller risk) in each cell. Soundness and monotonicity therefore require the two adjacent cells to have the same qualitative label. Iterating, all cells must have the same label, contradicting the monotonicity assumption that f(**0**) = 0, f(**1**) = H. Thus:

COROLLARY: There is no sound, monotonic, direct qualitative rating representation of any quantitative risk function that is a continuous, increasing function of its arguments.

Although this result is ostensibly more general than Theorem 1, it only points out that a discrete step function cannot give a completely accurate representation of any continuous increasing function. By contrast, the proof of Theorem 1 shows that qualitative ratings can also make "large" errors (e.g., by rating a risk of size xY as qualitatively no higher than a risk of size xy, even if Y is quantitatively much greater than y.) Moreover, such errors can occur fairly frequently: when the unit square is partitioned into M x M squares by M qualitative labels for each component, for example, then the reversal probabilities can be as high as about 18% (plus or minus 1.5% as M varies from 1 to 10, respectively) when a worst-case distribution, corresponding to the maximum values of ε, is considered according to Figure A1 and the proof of Theorem 1 above, although they are only about 4% for a bivariate uniform distribution with M = 3 as obtained by computer simulation.

This page intentionally blank

Chapter 2

Risk Analysis: Goals and Methods

1. INTRODUCTION TO HEALTH RISK ANALYSIS

Health risk analysis quantifies probable human health consequences, both positive and negative, of alternative risk management actions. It provides methods, principles, and high-level procedures for using scientific data to assess and compare the probable human health consequences of different exposures to hazards (i.e., sources of risk); to assess the likely changes in exposures and risks arising from alternative risk management decisions or interventions; and to evaluate and choose among alternative risk management interventions based on their probable health consequences. Health risk analysis is often divided into the stages of *risk assessment, risk management,* and *risk communication,* organized as an iterative process. Table 1 summarizes several traditionally defined steps in this process.

1.1 Risk Assessment

The first stage, health risk *assessment* estimates the health risks to individuals, to the entire population, and to selected subpopulations (e.g., infants, the elderly, immunocompromised patients, and so forth) caused by hazardous exposures and by the decisions and activities that create them. Health risks of sporadic illnesses due to exposures to bacteria are defined as the changes in the frequencies and severities of adverse health effects caused by the exposures.

Table 1: Traditional Steps in Health Risk Analysis

Step	Purpose and Description	Relevant information and techniques
Hazard identification	Identify potential sources of harm or loss. These sources are called *hazards.* Hazard identification identifies possible adverse health effects of activities or exposures and possible causes of observed adverse effects.	• Human data: Epidemiology, clinical and public health statistics; surveillance data. • Animal tests and bioassays • *In vitro* tests • Structure-activity patterns, molecular modeling, pattern recognition and statistical classification techniques
Exposure assessment	Quantify the number of people receiving various levels or intensities of exposure to a hazard over time. Relevant exposure metrics may depend on dose-response relations.	Environmental fate and transport models, possibly summed over multiple media (paths) and sources Studies of human activity patterns Biological monitoring of exposed individuals and receptors
Quantitative exposure-response and dose-response modeling	Quantify the magnitude of risk created by exposure of a target to a hazard. Characterize the probable frequency and severity of adverse health outcomes or losses caused by exposure to the hazard.	A quantitative risk assessment (QRA) runs multiple exposure scenarios through *dose-response models* to predict likely health impacts. Statistical, simulation, or biomathematical models of biological processes are used to quantify dose-response relations.
Risk characterization and uncertainty analysis	Combine estimated probabilities and severities of health harm (adverse consequences), together with indications of uncertainty or confidence, to create an overall summary and presentation of risk.	Monte Carlo simulation calculates risks by sampling many scenarios. Risk profiles, probability distributions, and trade-off and sensitivity analyses display risk, uncertainty, and variability.
Risk communication	Deals with how to present risk information to stakeholders. Considers how different types of recipients perceive risks and internalize/act on messages about them, in deciding what messages to send via what media.	Psychological theories and models and behavioral/ experimental findings on risk perception and effective risk communication.
Risk management decision-making	Decide what actions to take to control risks and hazards – i.e., accept, ban, abate, monitor, further research, reduce, transfer, share, mitigate, or compensate.	Risk-cost-benefit analysis, formal decision analysis for groups and individuals, risk quantification and comparison

Quantitative Definition of Risk

For sporadic illnesses, as opposed to epidemics, individual and population health risks can be defined more specifically as follows:

- The *individual risk* of sporadic illnesses caused by an exposure is the *expected number and severity of additional adverse health effects per capita-year caused by that exposure*. It can be tabulated or plotted as the expected number of cases per year in each severity category (e.g., mild, moderate, severe, or fatal, as defined in Buzby, *et al.*, 1996 based on illness-days and mortality). To avoid having to carefully define, describe and compare the severities of different illnesses, one may simplify by using illness-days per year for each type of illness (e.g., mild, moderate, or severe diarrhea) to summarize morbidity impacts. Both morbidity and mortality probabilities can also be given for different age groups or other population sub-groups. Alternatively, and often more conveniently, the expected loss of quality-adjusted life-years (QALYs) per case due to increased mortality and morbidity can be used as a single summary measure of severity, if the preference assumptions justifying QALYs are accepted (Hazen, 2003; Miyamoto, 1999). Individual risk is then given by the joint probability distribution of the number of cases per capita-year and the severities (i.e., QALYs lost per case) of these cases.
- *Population risks* are expressed as *expected numbers of additional adverse health effects per year* of each type or clinical severity category occurring in the population. They are the sum of individual risks over all person-years in the population. Population risks may also be further described by identifying subpopulations with especially high individual risks from exposure.

Technical Note: Use of Expected Values Use of the expected number of events per year to quantify risk is justified for sporadic illnesses that occur independently, or with only weak statistical dependence, in large populations, when the Poisson approximation (Janson, 1994) or compound Poisson approximation (Barbour, 2000) hold. The expected number of cases per year then determines the full probability distribution of the number of illnesses per year, to a close approximation (made precise in these references). Moreover, the Poisson probability distribution is stochastically increasing in its mean; thus, more expected cases correspond to less preferred distributions for all decision-makers who prefer fewer cases per year to more. If total illness-days per year result from a random number of cases, N, each independently creating a random number of illness-days, Q, then the total number of illness-days expected is $E(N) \times E(Q)$, independent of the specific probability distributions of the random variables N and Q. In this context, for a fixed distribution of Q, population risk can be interpreted as being proportional to E(N).

The formulas *Individual risk = E(illnesses per year) × E(QALYs lost per illness)* and *Population risk = sum of individual risks* are useful for sporadic illnesses, although they must be generalized for other types of risks, e.g., to allow for risk aversion (Cox, 2001). (If QALYs lost per year in the population are the sole quantity of concern, its mean is E(N)E(Q) and its variance is E(N)Var(Q) + Var(N)E^2(Q) (Feller, 1968). For Poisson-distributed N, E(N) = Var(N), and total QALYs lost per year in a large population will be approximately normally distributed with mean μ = E(N)E(Q) and variance σ^2 = E(N)[E^2(Q) + Var(Q)]. Mean and variance increase with E(N), so that, regardless of how they are combined to form certainty-equivalents, smaller values of E(N) will be preferred by any decision maker who prefers smaller means and variances of N to larger ones.)

Example: Tabulating Population Risk by Age and Sex

The following table, reproduced from Christensen *et al.*, 2001, shows the empirical estimates of population risk of campylobacteriosis in Denmark for 1999, broken down by age group and gender.

Incidence of Infections with Campylobacter, by Age and Sex in Denmark, 1999

Age group (years)	Number of cases				Cases per 100.000		
	Female	Male	Unknown	Total	Female	Male	Total
<1	33	39	6	78	103	115	118
1-4	180	258	42	480	133	181	172
5-9	63	118	21	202	39	69	61
10-19	190	218	37	445	67	74	77
20-29	584	419	117	1120	162	113	153
30-39	277	341	79	697	69	82	85
40-49	188	190	35	413	51	51	56
50-50	169	165	18	352	47	45	49
>60	156	181	40	377	26	40	36
Total	1840	1929	395	4164	68	73	78

Source: Christensen *et al.*, 2001
http://www.foodriskclearinghouse.umd.edu/poultry_campylobacter.cfm

In most age groups, men have a higher average risk (cases per 100,000 capita-year) than women. The total population risk of 78 cases per 100,000 capita-year (a size-weighted average of the rates in the different age- and gender-specific groups) is considerably higher that the reported rate in the United States of about 13.4 cases per 100,000 capita-year (CDC, 2003). Infants and young children have relatively high risks, and young adults (20-29) have anomalously high risks, perhaps because of changes in eating habits or exposures (e.g., home-cooked *vs.* other meals) and/or kitchen hygiene (Altekruse *et al.*, 1999). For both males and females, risk decreases

with age after age 29. Whether this reflects acquisition of immunity, decreases in exposure, or other factors is not yet known.

Example: Individual and Population Risks Caused by an Exposure

Problem: Suppose that in a population of 50 men and 50 women, 100% of the men and 20% of the women regularly eat a certain food (e.g., a raw or undercooked meat or dairy product) that exposes them to low levels of a bacterial pathogen. If women are not vulnerable to infection from this low-level source (i.e., risk from this food for exposed women = 0), and if men acquire immunity from it that protects them against larger exposures from other sources (e.g., in untreated milk or drinking water), thus cutting the expected number of illness cases per man-year due to this type of bacterium from 1 to 0.5, then what are the individual and population risks caused by eating this food? Is exposure positively or negatively associated with risk in this population? What would be the public health consequence in this population of reducing exposure to the food to zero?

Solution: For women, the food has no effect, and therefore the individual risks for women are zero. For men, the effect of the food is to reduce the expected illness cases per year from 1 to 0.5, for a total reduction in individual risk of 0.5 expected cases prevented per man-year. For the entire population, the reduction in risk for a randomly selected individual is Pr(woman) × (risk reduction for women) + Pr(man) × (average risk reduction for men) = 0.5 × 0 + 0.5 × 0.5 = 0.25 expected cases prevented per capita-year.

Despite these beneficial causal effects, exposure is statistically associated with increased risk. The risk to a randomly selected exposed person (who has 5/6 probability of being a man, since 50 men and only 10 women are exposed) is Pr(woman | exposed) × (risk for woman) + Pr(man | exposed) × (risk for man) = (1/6) × 0 + (5/6) × (0.25) = 0.21. But the risk to an unexposed person (who must be a woman) is 0. Thus, exposure is *statistically associated* with an increase in risk from 0 to 0.21, although it *causes* a reduction in risk from 1 to 0.5 for men (and has no effect for women.) For risk assessment and risk management, it is the causal effect of exposure, rather that the statistical association of risk with exposure, that should be used to quantify impacts of interventions that change exposure. Thus, for example, the impact of reducing exposure to zero would *not* be to prevent (0.21 excess cases per exposed individual) × (60 exposed individuals) = 13 expected cases prevented per year, as a naïve causal interpretation of the statistical relation between exposure and risk might suggest. Instead, it would be to *increase* risk in the population (and, specifically, in the subpopulation of men) by (50 men) × (0.5 excess cases per man-year) = 25 expected additional cases per year.

As shown in Table 1, following a National Academy of Sciences framework for risk analysis (Jaykus, 1996), the US Food and Drug

Administration (FDA), Centers for Disease Control and Prevention (CDC) and US Department of Agriculture (USDA) have defined risk assessment as a process that "consists of the following steps: hazard identification, exposure assessment, hazard characterization (dose-response), and risk characterization" (http://www.foodsafety.gov/~dms/lmriskgl.html). Dose-response assessment, in turn, consists of "The determination of the relationship between the magnitude of exposure and the magnitude and/or frequency of adverse effects." Similar concepts have been adopted internationally in WHO/FAO guidelines and OIE guidelines.

Chapters 3 through 5 discusses risk assessment more fully. The main goal of risk assessment is to produce information to improve risk management decisions. It does so by *identifying and quantifying valid cause-effect relations between alternative risk management decisions and their probable total human health consequences* and by identifying decisions that make preferred outcomes more likely. Health risk assessments typically use explicit – and, if possible, validated – analytic models (e.g., statistical, biomathematical, or simulation models) of causal relations between actions and their probable health effects. In general, quantitative risk assessment applies specialized models and methods to quantify likely exposures and the frequencies and severities of their resulting health consequences.

1.2 Risk Management

Health risk management (Section 6 of this chapter, page 64) applies decision analysis principles and other principles of rational choice to help identify and choose among alternative policies or actions that affect exposures, health risks, or their consequences. Risk management is often viewed as a process that takes scientific information obtained from risk assessment as an input, along with value judgments and policy goals and constraints, and that recommends choices of risk management actions as output. Alternative risk management approaches may include risk acceptance, prevention or avoidance (e.g., by reduction of microbial loads during processing or food preparation), mitigation of consequences (e.g., by appropriate clinical screening, diagnosis, and prescription procedures), transfer (e.g., health insurance), or compensation.

1.3 Risk Communication

Health risk *communication* (Section 7 of this chapter) characterizes and presents information about health risks and uncertainties to decision-makers and stakeholders. Risk assessment and risk communication should support effective risk management decision-making by providing the scientific information needed to compare alternative risk management

interventions in terms of their probable impacts on exposures and resulting changes in the frequency and severity of adverse health effects. For example, if animal antibiotics reduce the frequency and severity of some adverse human health effects, then these impacts should be included in the complete risk assessment and communication package and should be taken into account in risk management decision-making.

2. PURPOSES AND OUTPUTS OF RISK ANALYSIS

The primary purpose of health risk analysis is to support improved risk management decision-making. By definition, "better" risk management decisions are those that are more likely to produce preferred consequences, i.e., fewer illnesses, mortalities, illness-days, and treatment failures per person-year. Health risk analysis helps to identify such decisions, given whatever information is available when the decisions must be made. Health risk analysis also provides a framework for rational deliberation, information-seeking, conflict resolution, policy-making, and international and inter-agency harmonization about human health risks of commercial activities. When well conducted, health risk analysis can allow better-informed and more effective regulation of the production, distribution, preparation, and use of antimicrobials in food animals than approaches that are not driven by analysis of probable consequences of alternative decisions. Risk analysis models can predict how such activities interact with human behaviors – e.g., consumer or food worker behaviors in food handling and kitchen hygiene; physician decisions about what tests and treatments to recommend to which patients; and patient decisions about seeking and complying with physician instructions on antibiotic use – in determining the frequencies and magnitudes of adverse health outcomes.

The risk management decision alternatives to be evaluated by risk assessment typically include the following types:

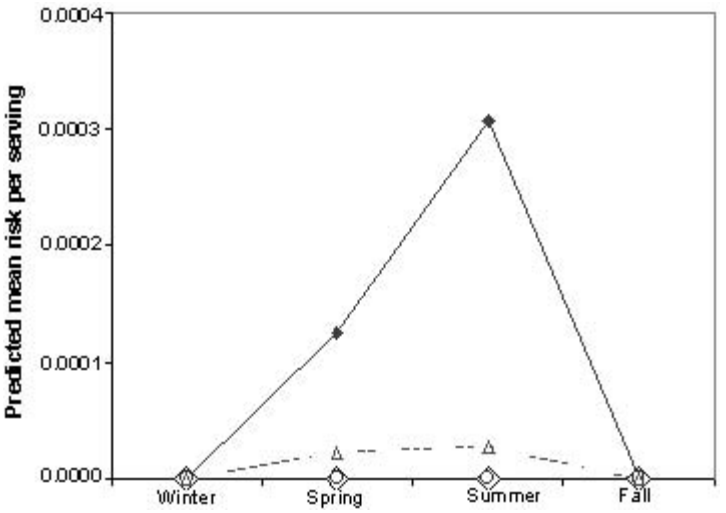
- *Status quo option:* Do not take actions to change current exposures.
- *Restriction or ban:* Intervene to reduce current exposures to hazards. Examples of interventions may also include training and education programs, monitoring and enforcement activities, and Hazard Analysis and Critical Control Point (HACCP) programs (USDA/FDA, 2004), as well as prudent use practices for antibiotics on farms and bans or restrictions on current uses of antibiotics.
- *Approval of new product or process:* Take action that may modify current exposure patterns (e.g., approve a new animal antibiotic product, use, or product line extension).

Different data are typically available for evaluating these three types of options. Changes in uses of products that have already been used for many years typically have the most data. Approvals of new products may have to rely more on plausible worst-case assumptions, models, and/or analogies to existing products to identify and bound potential risks. However, the same general logical assessment process applies to all risk management options. It is worth emphasizing that risk assessment intended to support improved decision-making should assess the changes in risks created by *alternative decision options or actions*, rather than the risks from the *status quo* alone or from specific products, classes of antibiotics, or situations. In other words, rational choice requires comparing alternatives.

Example: Comparing Predicted Effects of Alternative Interventions

Figure 1 shows model-predicted risks from eating Gulf Coast oysters, for several potential risk management interventions (*status quo*, rapid cooling, heating, freezing), by season of the year. These curves were produced via Monte Carlo simulation using a microbial risk assessment model detailed in FDA-CFSAN, 2001. The predicted changes in corresponding population risks are proportional to these changes in risk-per-serving, multiplied by servings consumed per season, and summed over all four seasons.

Figure 1: Predicted effects of alternative mitigations on mean illnesses per serving from *V. parahaemolyticus* (Vp) in Gulf Coast oysters.



Key: Top curve = no intervention, triangles = rapid cooling, circles = heat treatment, open diamonds = freezing

Source: FDA-CFSAN, 2001. http://vm.cfsan.fda.gov/~dms/vprisk6.html

As explained in the FDA-CFSAN report, "The effect of three Post Harvest mitigations was evaluated in the simulation: (a) mild heat treatment (5 min at 50°C), (b) freezing (-30°C), and (c) rapid cooling immediately following harvest (e.g., aboard ship). … The effect of mild heat treatment has been shown to reduce the density of *V. parahaemolyticus* to nondetectable levels (at least a 4.5 $\log_{10}$ reduction) and freezing at -30°C has been shown to reduce the density by approximately 2 logs. All three potential mitigation strategies have a substantial effect on the distribution of probable number of illnesses. The effect of these mitigations was evaluated under the assumption of the Beta-Poisson dose-response model [see Chapter 5, Figure 2]. For the Gulf Coast summer harvest, a shift in the distribution of probable number of illnesses down from a mean of 3,000 illnesses to approximately 240 illnesses is predicted under the mitigation of rapid cooling. The mean number of illnesses projected to occur under the freezing mitigation is approximately 15. …The simulation results suggest that in the absence of subsequent post harvest mitigations, 'at harvest' guidance levels of 5 log (10^5), 3 log (10^3) and 2 log (10^2) total *V. parahaemolyticus* per g could (potentially) reduce the illness rate by 2%, 50% and 90% with corresponding losses of 0.3%, 25% and 70% of the harvest, respectively."

Example: Assessing Risk Impacts of a Non-Specific Intervention

Even if specific interventions have not been identified, a risk model can be used to show the potential gains in public health from interventions that reduce exposures by stated amounts. For example, the following table shows the model-predicted change in population risk caused by an intervention that reduces *Salmonella enterica* (non-typhoid *Salmonella*) concentrations in servings of broiler chicken by 50%. How such a reduction might be achieved in practice is left unspecified. The value of the model is to show how large an effect on public health such an intervention would have – a reduction in expected illnesses per 100,000 capita-years from 29 to 11 – even before any detailed means for achieving it have been proposed. (For details of this risk assessment model, see WHO/FAO, 2002.)

Summary of risk before and after a 50% reduction in *Salmonella* concentrations

	Before	**After Intervention**
Prevalence	20%	20%
Expected risk per serving	1.13E-05	4.28E-06
Number of servings in year	26	26
Annual expected risk	2.94E-04	1.11E-04
Rate of illness per 100 000	29	11

Source: WHO/FAO, 2002

Example: Decision-Relevant Scope for a Risk Assessment

Suppose that a regulatory agency publishes a risk assessment of an animal antibiotic, A, that it suspects might cause resistance in bacteria B. These bacteria can be transferred to humans via food, where they may cause A-resistant infections. Antibiotic A is also used in human medicine to treat foodborne infections caused by bacteria of type B, so A-resistant strains of B are of concern.

The agency states that "This risk assessment seeks to inform decision-making about whether to permit continued use of antibiotic A in food animals, by providing an estimate of the number of cases of human bacteremias per year caused by bacterium B that are resistant to antibiotic A, where the resistance is potentially linked to food animal uses of A." Is this an appropriate scope for the risk assessment of animal antibiotic use A? Why or why not?

Solution: The announced scope can be improved in the following respects:

1. *It identifies only one decision option* to be assessed, namely the *status quo:* continuation of "food animal uses of A". It does not identify or assess any alternative actions (e.g., discontinuing food animal uses of A) for comparison. Thus, the results will not provide decision-makers with the information needed to compare and choose among competing options.
2. *This scope does not address human health harm.* For example, suppose that resistance has *no effect* on human health or on treatment efficacy. (This can happen in practice when "resistance" is defined via an elevated mean inhibitory concentration (MIC) for the antibiotic in *in vitro* tests, but therapeutic doses are high enough to kill even the resistant bacteria.) Then, logically, the risk attributed to resistance should be zero since, by assumption, it has no effects. Yet, the "number of cases of human bacteremias per year caused by bacterium B that are resistant to antibiotic A" might nonetheless be large. Thus, it is not the right quantity to estimate to understand risk. Instead, cases of failed or compromised treatment are the appropriate quantity.
3. The risk assessment should be based on the *change* in human health harm caused by bacterium B due to use of antibiotic A in food animals. The number of cases in which there is a change in human health harm is not the same as (and may be much smaller than) the total number of resistant cases. For example, if the adverse human health effect of concern is treatment failure brought about by resistance to drug A in bacteria infecting human patients, then cases in which the patient is not prescribed antibiotic A, or would not benefit from treatment with A for reasons not related to resistance (e.g., inability to tolerate A), cannot lead to additional treatment failures, and so should not be counted.
4. The number of cases that are "potentially linked to food animal uses of A" may be much larger than the number of cases actually caused by use of A. While the meaning of "potentially linked to" is not given, it might be interpreted as

> "attributable to" in the sense of epidemiological measures of attributable risk. These measures typically reflect statistical association, rather than causation. But effective risk analysis requires quantification of the causal relation between changes in exposures (or in the actions leading to them) and changes in probable human health consequences.

In summary, the stated scope seeks to estimate a quantity that is larger than the number of cases in which use of A in animals causes excess harm in humans. It does not compare the harm that would occur for alternative actions (e.g., banning *vs.* continuing animal use of A). These comparisons are needed to inform rational (consequence-driven) risk management decisions.

The main value and purpose of risk assessment is usually to quantify and compare the probable human health risks (i.e., changes in the expected number and/or severity of foodborne illness cases per year in exposed populations) for *each* risk management decision option considered, *conditioned* on whatever information is available about it. Computational-statistical, mathematical and probability modeling, and computer simulation methods enable risk assessors to estimate quantitative bounds on human health risks and uncertainties from realistic (incomplete, imprecise, inaccurate and perhaps inconsistent and incorrect) measurements and data.

3. RISK ANALYSIS WITH UNCERTAIN DATA

Risk analysis maps technical inputs, describing how decision alternatives affect the number of people exposed to a hazard and the likely adverse consequences of such exposures, into quantitative assessments of risk (e.g., expected illness cases per year and QALYs lost per case) for each alternative. In practice, many of these inputs are uncertain. For example, Table 2 summarizes data gaps identified in risk assessment for campylobacter in broiler chicken. Similar data gaps have been identified for other microbial risk assessments of *Listeria* in ready-to-eat foods, *Escherichia coli* in ground beef, *Salmonella* spp. in eggs and in broiler chickens, and *Vibrio parahaemolyticus* in fish and shellfish (RAC, 2004).

A pragmatic perspective on risk analysis with uncertain input data is that (a) Risk model inputs are almost always uncertain in practice; but (b) Risk management decisions can still be informed and improved by uncertain inputs (e.g., based on imperfect measurements and incomplete facts, knowledge and data) so long as they provide some statistical information about probable human health consequences of alternative decisions.

Table 2: Data Gaps in Risk Assessment of *Campylobacter*

Information Type	Specific Need
Dose-response	Additional data on dose-response.
Strain variability	Data on strain variability in relation to virulence and pathogenicity.
Strain variability	Data on strain variability in relation to survival during processing.
Virulence/ pathogenicity	Studies on the mechanisms of infectivity, virulence/pathogenicity of *Campylobacter* in the human host.
Virulence/ pathogenicity	Studies/data on the development of antimicrobial resistance; transference to human host
Pathogenicity	Data/studies on links to cancer in human host
Dose-response	Quantitative information about infection and illness rates at low doses of C. jejuni, and also at a range of doses of different strains of *C. jejuni* and strains of *C. coli*.
Epidemiology	Complete epidemiological data from outbreak studies including enumeration of thermophilic *Campylobacter* in suspected food items or in drinking water, numbers of people exposed, attack rates, and demographics of those exposed, particularly immunocompromised population groups and children under the age of five.
Epidemiology	Enhanced surveillance and outbreak investigations
Immunity	Data describing the impact of and longevity of acquired immunity resulting from recent exposure to thermophilic *Campylobacter*.
Exposure	
Survival	Information about the influence of strain-specific variation on *Campylobacter* survival on poultry meat.
Prevalence	Survey data on the prevalence of *Campylobacter*-positive flocks for slaughter, that includes information on sample size, test methods etc.
Contamination	Data on the routes of *Campylobacter* colonization of broilers at the farm level so that farm interventions can be appropriately targeted.
Contamination	Data on the probability of contamination of birds during transport.
Contamination	Studies on dynamics of within-flock transmission of *Campylobacter*.
Prevalence/ enumeration	Prevalence and enumeration data for *Campylobacter* on carcasses before and after various processing steps such as scalding, defeathering, evisceration, washing and chilling.
Prevalence/ enumeration	Prevalence and enumeration data for *Campylobacter* on carcasses comparing various methods of chilling (e.g. air chilling, water chilling, water chilling with chlorine).
Prevalence/ enumeration	Prevalence and enumeration data for *Campylobacter* on carcasses comparing different scalding temperatures or alternate scalding configurations (e.g. multi-tank scalding systems).
Contamination	Data describing actual cross-contamination between positive and negative flocks and within flocks during different slaughter processes.
Handling/ preparation	Additional data on the cooking of chicken that addresses areas of the chicken where *Campylobacter* may be protected from heat.
Handling/ preparation	Survey and direct observational data on consumer practices in preparing and handling chicken that detail frequency and degree that transfer and subsequent ingestion of *Campylobacter* could occur.

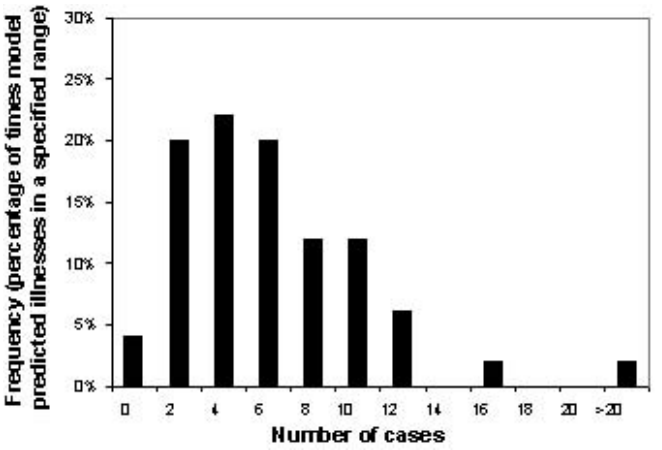
Source: Adapted from RAC, 2004

This perspective is sometimes formalized in "value-of-information" (VoI) calculations and sensitivity analysis calculations (Yokota and Thompson, 2004). It allows risk assessment to deliver useful results about probable risks and remaining uncertainties based on currently available empirical information, while also showing the potential for these results to change as further information is collected. A thorough analysis also indicates which new information is likely to lead to the greatest changes in current estimates. This information, in turn, is what risk management decision-makers need to make rational interim decisions and to identify what new empirical information would be required to justify future changes.

Example: Displaying an Uncertain Population Risk

Figure 2, from FDA-CFSAN, 2001, shows an estimated probability distribution for the number of septicemia cases per year in the United States caused by eating oysters contaminated with *V. parahaemolyticus* (Vp), assuming that no risk management interventions are undertaken. Its mean value is 6 cases per year. If this were a Poisson distribution, then the mean would fully determine the entire probability distribution, and it would be unnecessary to specify anything other than this mean value to fully characterize risk. But in reality, the probability distribution was generated via a Monte Carlo uncertainty analysis in which many input parameters and assumptions were sampled from input probability distributions, resulting risks were then calculated, and the process was repeated thousands of times to estimate the entire probability distribution of risk. (See FDA-CFSAN, 2001 for details.) Thus, in effect, there is *uncertainty* about the true risk.

Figure 2: (FDA-CFSAN, 2001) Model-predicted probability distribution of Vp-related septicemia cases per year in United States if no interventions are made



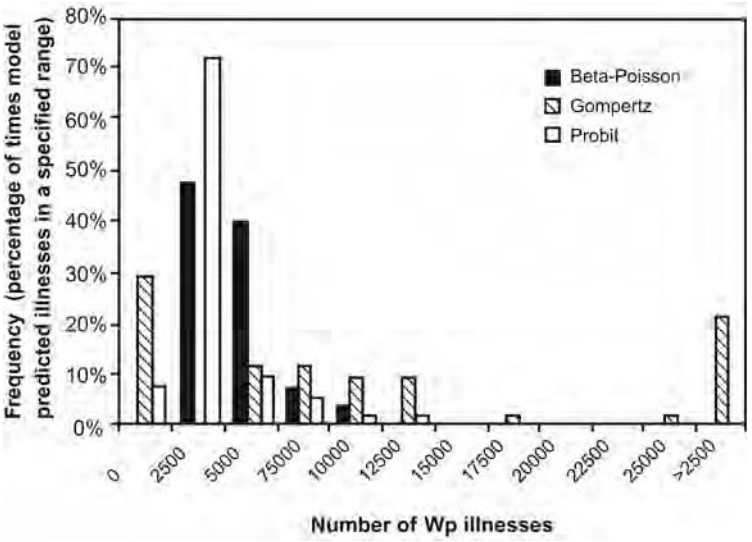
To interpret such a model-based probability distribution for risk, it is important to understand what assumptions have been made – and which of these assumptions are held fixed (i.e., "conditioned on") – in generating the displayed risk distribution. Sensitivity analysis plots (see Chapter 5, Section 5) can help to unravel the web of assumptions behind the distribution and indicate which assumptions most affect the uncertainty distribution of risk and how much higher or lower than the estimated mean value the true value in any year is likely to be. Before looking at the sensitivity analyses, one should only infer from a summary plot such as Figure 2 that the estimated number of cases per year predicted by the model (and by the assumptions and input distributions behind it) is likely to be in the range from 0 to 14 and relatively unlikely to exceed 20.

It is also important (and is now standard practice) to distinguish between *uncertainty* and *variability* in presenting distributions of risk. For example, some people eat more oysters than the average consumption level, increasing their risk of Vp-related illness. Others may be more susceptible than average (e.g., due to reduced immune system function), or live in areas with warmer water and air (where bacteria multiply more quickly). These factors create differences in the risk per capita-year for different people that cannot be eliminated by taking larger sample sizes or more accurate measurements. They result in true differences in the exposures and risks experienced by different individuals. Such heterogeneity in individual risks is called *variability* in the risk. It can be described by a population frequency distribution of values. Each individual has an expected number of illnesses per year based on his or her own exposures, attributes, behaviors, and other covariates, and the frequency distribution of risk in the population specifies the fraction of the population having at most each level of individual risk.

When both uncertainty and variability in risk are present, as is usually the case in practice, the expected number of cases per capita-year can be thought of as having a true but unknown frequency distribution in the population, with individual values that may depend on individual covariates such as age, sex, weight, immune status, and geographic location, as well as on exposure. Uncertainty about this true frequency distribution can be conceptualized as a set of probabilities for different frequency distributions (see Chapter 5). Even if the probabilities of different frequency distributions are uncertain, simply displaying the conditional frequency distributions based on alternative sets of assumptions can provide decision-makers with useful information about the plausible range of population risks. For example, Figure 3 shows the effects of assuming three different dose-response models on the predicted frequency distribution of risk (illnesses per year) from Louisiana Gulf Coast summer oysters. Clearly, the Gompertz dose-response model predicts a higher probability for large numbers of

illnesses (e.g., > 25,000 illnesses per year) than the other two dose-response models. This type of information can be useful for understanding the range of possibilities, even if Bayesian model-averaging (BMA) or other statistical methods for quantifying model uncertainties (Cox, 2001) have not been used to assess the relative probabilities of the different dose-response models.

Figure 3: Effect of dose-response model on predicted illnesses per year from *V. parahaemolyticus* (Vp) consumption in Louisiana Gulf Coast summer harvest



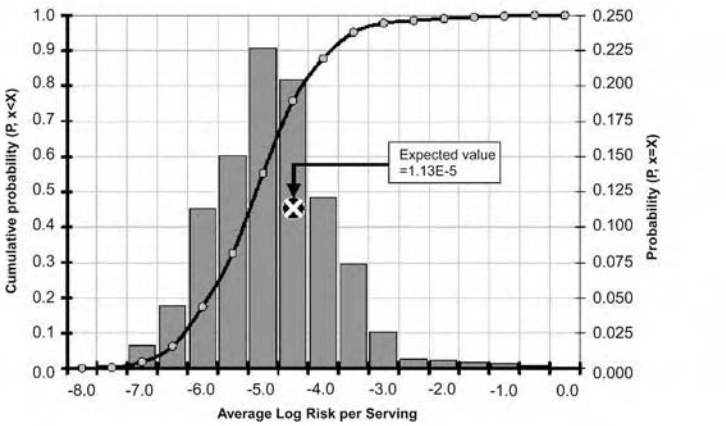
Source: FDA-CFSAN, 2001. http://vm.cfsan.fda.gov/~dms/vprisk6.html

Example: A Risk Distribution for *Salmonella* Risk in Broiler Chickens

Figure 4a illustrates a model-based prediction of the probability distribution (gray histogram) and corresponding cumulative probability distribution (s-shaped black curve) for average illness risk per serving of broiler chicken caused by contamination with *Salmonella enterica* (non-typhoid *Salmonella*.) This figure is described in the WHO/FAO, 2002 risk assessment report as follows: "Assuming a 20% prevalence of contaminated broilers, the estimated frequency and cumulative distribution of average risk per serving are shown in [the figure]. The expected risk per serving is 1.13E-5, or 1.13 illnesses per 100 000 servings. This value represents the average risk for all individuals in the population that consume servings of chicken that are stored, transported and prepared in the manner described in the model, and also accounts for the probabilities that the serving was from a chicken contaminated with *Salmonella*, and that the meal was undercooked. It should be

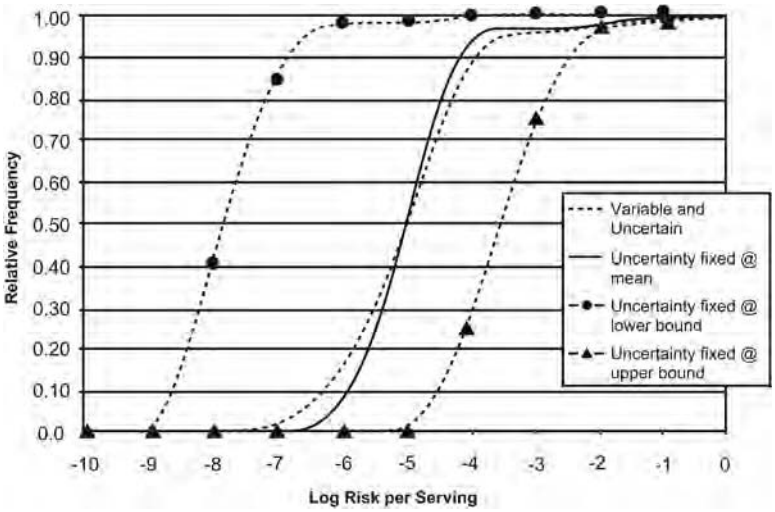
recognized that some individuals consuming a serving on certain occasions would experience a much higher risk than others who might be consuming servings with no salmonellosis risk at all, since the serving would be free of the pathogen."

Figure 4a: Distribution of Average *Salmonella* Risk per Serving of Chicken



Source: WHO/FAO, 2002

Figure 4b. Effects of parameter uncertainties on per-serving risk distribution.



Source: WHO/FAO, 2002

Figure 4b shows the sensitivity of the probability distribution of average risk in Figure 4a to several sources of uncertainty. The explanation in the WHO/FAO report is as follows:

"Several of the parameters in the cooking module were considered uncertain and are listed in [the following table]:

Uncertain parameters in the cooking module

Consumption relationship	Mean	Min.	Max.
Probability [that a serving is] not adequately cooked	0.1000	0.0500	0.1500
Proportion [of bacteria] in protected area [of chicken carcass, where they may survive cooking]	0.1567	0.1000	0.2000
Exposure time to cooking temperature of cells in protected areas	1.00	1.50	0.50
Cooking temperature reached in protected areas.	63.50	65.00	60.00

The impact of uncertainty in these parameters was investigated in order to evaluate their influence on the risk estimate. To do this, the model was re-simulated using a fixed single value for each of the uncertain parameters while allowing the other parameters of the model to vary within their defined distributions. Three simulations were performed: in the first, the parameters listed in [the above table] were set at their mean value[s]. The fixed values used for the second simulation were those that would generate a "worst case" scenario, i.e. the maximum value for probability that the chicken was undercooked, the maximum value for proportion of cells in a protected region, the minimum heat exposure time, and the minimum value for the temperature reached in a protected region (0.15, 0.2, 0.5 minutes and 60°C, respectively). It is recognized that such a scenario may not occur in reality, but it gives an upper bound to the range of possible values. The third simulation used the values that would give a "best case" scenario, i.e. minimum value for probability undercooked, etc. This approach allowed the extremes in the risk distribution, driven by the uncertain parameters, to be highlighted. The results of performing the analysis on the uncertain parameters influencing consumption risk are shown in [Figure 4b].

When the uncertain parameters were fixed at their mean values (Uncertainty fixed @ mean) and compared with the risk distribution generated by the model when all parameters were allowed to vary (Variable and Uncertain), it appears that within the range of uncertainty that was assumed to define the parameters, the impact of variation is not very large. The resulting risk distributions are similar and the tails of the currently defined uncertainty distributions do not have a dramatic impact on the overall risk uncertainty distribution. In other words, the range and shape of the distributions defining uncertainty do not influence the risk uncertainty significantly.

> Alternatively, if the assumptions made were incorrect and the uncertain parameters actually spanned a different range, e.g. if the true values are centered nearer to the min. or max. values rather than at the value assumed to be the mean, the distribution of risk would approach the extreme distributions shown. … A complete quantitative uncertainty analysis of the model and all input parameters was beyond the scope of this work. This type of analysis is time consuming and not necessarily more informative for the purposes of this document. Many of the inputs are generic approximations in order to provide a representative risk scenario. Nevertheless, it is important to recognize these two characteristics – uncertainty and variability – in the probability distributions used in quantitative risk assessments. It is also readily recognizable that several input parameters in this model are *both* variable and uncertain, and, if the individual parameters are important in determining the magnitude of the risk estimate, it may be necessary to separate the uncertainty and variability in the quantitative analysis in order to understand their impacts and arrive at proper risk estimations." (WHO/FAO, 2002)

In summary, although the distribution in Figure 4a conflates uncertainty and variability, the sensitivity analysis in Figure 4b helps to unravel their separate contributions, indicating the extent to which further information might eventually change the estimated risk per serving.

4. DESIRED OUTPUTS OF RISK ANALYSIS

A successful risk analysis shows the estimated changes in the frequencies and magnitudes of adverse human heath consequences resulting from different risk management decision options. (Of course, if hazard identification and risk management reveal that the risk from the *status quo* is so small that no risk management action is needed, then risk analysis may stop there. A full risk analysis is usually carried out when a risk management intervention is being contemplated, and we will henceforth assume that this is the case.) Risk analysis uses probability distributions, confidence intervals, and other displays to show uncertainties about the human health consequences of different decisions. It identifies a subset of one or more decision options leading to preferred probability distributions of health risks and other outcomes. Thus, a successfully completed risk analysis should allow a decision-maker to answer the following questions for each risk management decision alternative being evaluated or compared:

- *What probable change in human health risk would result from each risk management intervention?* If the risk management decision option or action being assessed is implemented, how will the probable adverse human health effects (e.g., expected numbers of mild, moderate, severe,

and fatal illnesses per year; expected numbers of illness-days and, if desired, quality-adjusted life-years (QALYs) lost per year) change in the whole population and in subpopulations with distinct risks?

- *How certain is the change in human health risk that would be caused by each risk management action?* Instead of a single value, i.e., a "point estimate" of risk, uncertain risks are characterized by intervals or probability distributions indicating how closely the change in human health risk caused by a proposed risk management intervention can be predicted. There are several technical options for expressing uncertainty around point estimates (e.g., plausible upper and lower bounds or confidence limits, coefficients of variation, tolerance intervals, prediction intervals, Bayesian probability intervals, Bayesian posterior distributions, etc.) More elaborate uncertainty displays (e.g., confidence contours for the joint distribution of frequency and severity components of risk) are available for specialists. The essential information to provide about uncertainty in any risk assessment is how large or how small the true risks might be, consistent with the data and with the specified assumptions of the risk assessment. Point estimates that are "best" with respect to various technical statistical criteria will typically fall between these extremes.

  Technical note: Statistical point estimates and interval estimates. Many criteria have been used to define and identify "best" point estimates in risk models, e.g., maximum likelihood estimates (MLE), maximum *a posteriori* (MAP) Bayesian estimates, maximum entropy, minimum description length, least squares, minimum absolute deviation, and minimum expected loss (for various loss functions) (see Cox, 2001 for a survey of methods for risk analysts). While these criteria have led to useful theory and algorithms for estimating the parameters of risk models, *none* of them is satisfactory as the sole output from a risk assessment. *It is essential to provide intervals or probability distributions around any point estimate of risk* to inform the users of a risk assessment about the full range of risks that might be caused by a risk management intervention. This principle applies to qualitative and fuzzy risk ratings as well. If a point estimate of a risk is "High", then some indication must be given of how certain this value is and of how compatible the frequency and severity components of the risk are with other qualitative labels, such as "Low". A risk assessment that produces a single overall value for risk with no indication of uncertainty should be avoided.

- *What are the key drivers of risks and uncertainties for each option?* The analysis should make clear to the user the main reasons *why* the estimated risk from each decision option is as high or low as it is. Are the results driven mainly by predicted exposure levels, by the responses of sensitive subpopulations, by genetic or epidemiological data that

establish tight constraints on the plausible values, or by other factors? Sensitivity analyses that plot how estimated risks would change as input assumptions and estimates vary within plausible ranges (e.g., within a few standard deviations of their median values) can help to identify the combinations of input values that drive the main conclusions and the extent to which these could be changed without changing the comparison of different risk management interventions. Chapter 5 provides examples of such sensitivity analysis displays.

- *Which risk management interventions are undominated?* One risk management intervention *dominates* another if it produces smaller probabilities of exceeding any specified level of adverse consequences per year. For example, if two different interventions lead to different expected numbers of sporadic salmonellosis cases per year (with the actual number being a Poisson random variable), and if the probable health consequences per case (e.g., the distribution of the number of days of illness of given severity) is the same for each intervention, then the one giving the smaller expected number of illnesses per year dominates the other. If the expected number of cases per year for each intervention is uncertain, and if the probability that it exceeds any specified value is smaller for intervention A than for intervention B (for *all* possible specified values), then A dominates B. Scientific risk assessment can, at most, identify undominated risk management alternatives for risk managers to further assess and choose among, but stops short of being able to recommend an objectively "best" choice among multiple undominated interventions.

5. INTRODUCTION TO RISK ASSESSMENT

Risk assessment is the part of risk analysis that uses facts and data to predict the probable consequences of alternative risk management decisions. It builds and validates predictive causal models of risk, and uses them to predict how different interventions will change the probable frequencies and severities of adverse health consequences in a population. This section illustrates some key ideas of quantitative risk assessment, while Chapters 3 through 5, respectively, provide a more thorough treatment of hazard identification, exposure assessment, and risk characterization.

5.1 A Rapid Risk Rating Technique (RRRT)

Although risk assessment models often use technically sophisticated statistical, probabilistic, and simulation-based analytic methods to estimate

and validate the probabilistic input-output relations between actions and exposures and between exposures and health effects, the main logic of most health risk assessments is straightforward. For example, as explained in Chapter 1, the risk of a foodborne illness in a population can often be modeled as a product of a few high-level factors, as follows:

Population Risk = Expected adverse human health consequences per year in the population = (expected number of contaminated units, e.g., servings or bacteria, depending on the dose-response relation, ingested per year) × (expected illnesses per contaminated unit ingested) × (expected illness-days or other adverse consequences per illness).

This can be abbreviated, as in Chapter 1, as:

Risk = exposure factor × dose-response factor × consequence factor.

Such a multiplicative, "top-down" approach, meaning an approach that starts with estimates of high-level factors instead of with the more detailed network of low-level factors that determine them, has been recommended on methodological grounds (Bailar and Travers 2002; FSRC, 2003) and has long been used by risk assessment practitioners (e.g., McNab and Alves, 2003). While detailed review of past studies and data and modeling calculations are usually required to estimate and document the estimated values and uncertainty intervals (or approximate probability distributions) for these factors, the overall multiplicative logic is relatively simple and transparent. It can be applied to many risks caused directly by foodborne pathogens. Any intervention that changes one or more of these factors will change the predicted population risk correspondingly.

If a risk management intervention simultaneously affects multiple contaminants (e.g., multiple pathogens, or both susceptible and resistant strains of a pathogen), multiple food commodities, and/or multiple subpopulations having distinct exposure-response relations, then summing the above product over all combinations of these multiple components gives the total impact on population risk. This is the basis of the *Rapid Risk Rating Technique (RRRT)*, introduced here and discussed and developed further in Chapters 6 and 8. The RRRT approach expresses the risk of interest as a sum of products of factors. Each factor is estimated from data. Uncertainties in the estimates are expressed via intervals, bounds, uncertainty factors, or probability distributions. Each product of factors corresponds to a causal path leading from risk management actions, and/or changes in exposures that they cause, to resulting changes in health effects in a population. The sum over many such products allows the impacts for

different combinations of pathogens, food commodities, subpopulations, and so forth caused by a risk management intervention to be totaled.

Table 3 shows the beginning of a risk assessment calculation made in the RRRT framework. The example shown is estimation of the likely human health consequences of banning macrolides, a class of antibiotics used in both veterinary and human medicine, from further use in animal feed. (This example is developed further in Chapter 8.)

Table 3: Example Risk Assessment for Macrolide Use in Chickens

Variable	Values, Uncertainty Factors (UF)	Data Sources
Total current campylo-bacteriosis cases reported per 100,000 people/yr.	13.37 cases/100,000 in 2002 for FoodNet surveillance sample in US, UF ≈ 1 (i.e., little uncertainty)	CDC, 2003
United States population	292E6 = 2,920 × 100,000 people in US, UF ≈ 1	US Census Bureau
Fraction of *C. jejuni* cases severe enough to warrant antibiotic treatment	0.00595 (Uncertainty about this factor is analyzed via sensitivity analysis.)	Buzby, *et al.*, 1996
Average total severe cases per reported severe case	8 (Ranges from 2 for severe cases to 38 for mild cases; UF ≈ **5**)	Mead *et al.* (1999)
Fraction of severe cases that are *C. jejuni*	0.99 (May be as low as 0.95), UF ≈ 1, i.e., there is little uncertainty	CDC DBMD
Fraction of severe *C. jejuni* cases caused by chicken products (or via cross-contamination of other foods)	0.10, uncertainty factor = 3-**10**, estimate based on competing risk, genetic, epidemiological, and historical data, as discussed further in Section 4 of Chapter 4.	Chapter 4; Stern and Robach, 2003
Fraction of chicken-caused severe cases that are antibiotic-resistant	0.01 for erythromycin resistance, UF = **2**	CDC, 2000
Resistant severe *C. jejuni* cases per year = product of above	**1.84** cases/yr. for macrolides = (13.37E-5)*0.00595*8*292E6*0.99* 0.10*0.01; UF = 18 (from component UFs of 5, 10, 2)	Product of above.

The portion of the calculation shown estimates the expected number of *C. jejuni* cases per year that are: (a) severe enough so that they might potentially benefit from antibiotic treatment; but (b) macrolide-resistant; and (c) caused by consumption of chicken meat (or other food products that have been cross-contaminated by contaminated chicken). Like all RRRT calculations, this one is organized as a product of factors. The calculations use population average values of several of the factors, and no summation over multiple combinations of risk factors is performed to more fully describe risks to subpopulations. The bottom-line number in Table 3 is the

product of the factors above it. As described momentarily, uncertainties are also assessed using multiplicative "uncertainty factors", one of several options for uncertainty analysis of products.

Logically, the correct value of the total number of macrolide-resistant severe *C. jejuni* cases per year from chicken, estimated in this table as 1.84 cases per year, puts an upper bound on the number of such cases that could experience harm (e.g., more illness-days) due to resistance-related macrolide treatment failures that might be prevented by reducing macrolide-resistant *Campylobacter* in chicken products. It is an upper bound because not all such cases will actually be prescribed macrolides (e.g., some may be prescribed fluoroquinolones), nor will all cases that are prescribed macrolides necessarily result in lost health benefits. In addition, a ban on macrolide use in chickens probably would not eliminate all such resistant cases, at least in the foreseeable future. Nonetheless, estimating even a rough upper bound can provide useful information about the potential size of the human health benefits to be expected from a ban,

For purposes of illustrating risk assessment calculations, the most important points about the RRRT framework are as follows:

- *The scope of the assessment is matched to the specific risk management decision options being evaluated and compared:* The *scope* of the risk assessment calculation in Table 3 is to estimate the incremental number of macrolide-resistant *C. jejuni* cases per year that: (a) Might be caused by (or, more precisely, preventable by discontinuing) current use of current macrolide products in chickens; and (b) Are severe enough to warrant treatment with erythromycin or another human macrolide in current clinical practice, i.e., some clinical benefit might be achieved if the treatment is effective. Such a calculation can help bound potential human health benefits (risk reductions) from restricting or eliminating use of macrolides in chickens. It would not be appropriate for evaluating a ban on all macrolides used as growth promoters, nor for evaluating introduction of a new product (e.g., a macrolide product line extension). To complete the risk assessment, it is necessary to carry out similar calculations for other pathogens (e.g., macrolide-susceptible *C. jejuni*, and both susceptible and resistant *C. coli*) affected by the risk management intervention being assessed. Rather than pursuing this in detail, an initial rapid screening assessment might simply assume that risks from chicken-borne *C. coli* are not greater than those from chicken-borne *C. jejuni*, and use this assumption to bound the additional contribution from *C. coli*.
- *Transparent calculation logic* (Bailar and Travers, 2002). The calculations are based on multiplying factors estimated from documented data sources. Thus, if someone considers any of the cited

values to be inappropriate, or when more recent data become available, the specified values of the factors can be easily updated and the results recalculated. Although estimation of the values of the individual factors (and their corresponding uncertainty factors) from available data may involve a great deal of research and detailed statistical calculation, and the supporting arguments should be carefully documented and discussed in any full analysis, this does not obscure the simple multiplicative framework for combining the results to obtain a final risk estimate.

- *Clearly interpretable output.* The main result in Table 3 is that the number of "Resistant severe *C. jejuni* cases per year caused by chicken products" is estimated to be 1.84 cases per year. This point estimate is accompanied by an uncertainty factor (explained below) of about 18, corresponding to a subjective Bayesian 95% probability interval of [1.84/18, 1.84 × 18] = [0.1, 33] cases per year. In contrast to qualitative designations such as "high", "low", "acceptable", "unacceptable", and so forth, the result "1.84 cases per year" has a clear meaning that is directly useful for informing decisions.
- *Modular calculations.* The intermediate result of 1.84 cases per year does not yet consider the *consequence* component of risk, i.e., what fraction of these cases will seek medical care, be prescribed a macrolide antibiotic, experience treatment failure due to resistance, and suffer excess days of illness. Nor does it consider the *preventable fraction* of the exposures, i.e., the fraction of macrolide-resistant *C. jejuni*-contaminated chicken servings that would be removed (and presumably replaced by macrolide-susceptible *C. jejuni*-contaminated servings) in the event of a risk management intervention. By organizing the calculations as a multiplicative sequence, however, it becomes possible to stop the calculation part way through, yielding an upper-bound estimate of the final result of the complete exposure, illness, and consequence product calculation (since multiplication by additional fractions between 0 and 1 can only reduce the current result.) Thus, 1.84 is an upper bound in this sense (i.e., including additional factors will only reduce it further) for the point estimate of the preventable number of cases per year that may experience a loss of clinical benefits due to macrolide-resistant *C. jejuni* from macrolide-exposed chickens. (This should not be confused with a statistical upper confidence limit. Statistical uncertainty analysis is discussed below and in Chapter 5.)
- *Sensitivity analyses.* Sensitivity analysis is especially simple for product models. Inspection of the numerical values of the factors in the product shows which ones have greatest impact on the final results. In conjunction with uncertainty factors indicating how many times too high or low the estimated values might plausibly be compared to the true

values, the estimated point values show what changes in factors might occur as additional information is collected and by how much such changes could increase or decrease the current point estimate of risk. In Table 3, for example, the estimated fraction of cases that are severe enough to potentially benefit from antibiotic therapy is obviously a crucial parameter, as it reduces the overall product by a factor of 0.00595. This value is obtained from Buzby *et al.*, 1996. Increasing or decreasing it to reflect more recent data, when they become available, will increase or decrease estimated risks proportionally.

- *Uncertainty analysis.* Uncertainty factors of about 1 (UF ≈ 1) in Table 3 indicate quantities that are known with enough precision and confidence so that better information about them is not expected to make a large change in the results. Uncertainty factors greater than 1 indicate that the point estimate may plausibly be too high or too low by the amount of the uncertainty factor. This is only one way to indicate approximate uncertainties, but is often useful for multiplicative models. Combining the quantified uncertainty factors using a central limit theorem (discussed in the following technical note) gives an estimated uncertainty factor of 18 to the point estimate of 1.84 cases per year. This provides an indication of how many times larger or smaller than 1.84 the true but unknown rate of cases per year might plausibly be, based on the quantified uncertainties and point estimates in Table 3. In addition to these quantified uncertainties, as already mentioned, changing the 0.00595 point estimate of the fraction of severe cases (from Buzby *et al.*, 1996), or changing the scoping assumption that only these severe *C. jejuni* cases warrant treatment with antibiotics and might receive clinical benefits from such treatment, could lead to proportional changes in the point estimate of risk.

Technical Note: Simplified multiplicative uncertainty factors. To enable quick approximate uncertainty analysis without Monte Carlo simulation, it is convenient to make the artificial restriction that uncertainty about each parameter is approximated by a single multiplicative uncertainty factor. In other words, uncertainty about the point estimate x of an uncertain parameter X is expressed by an *uncertainty factor,* UF, such that the true value of X is considered equally likely to be above or below its point-estimated value, x, and there is a subjective 95% probability that the true value of X lies between x/UF and x × UF. The interval [x/UF, x × UF] is interpreted as a subjective Bayesian confidence interval for X. Although this simplified approach to uncertainty assessment is not flexible enough to represent arbitrary beliefs (e.g., it is inappropriate for representing quantities with zero as a plausible value, or proportions for which x*UF > 1), it does allow uncertainty about each model parameter to be expressed, at least approximately, by a single number.

Moreover, formulas for the human health benefits and risks from risk management interventions are often expressed as products of uncertain parameters, and thus may have approximately log-normal uncertainty distributions (Druzdzel, 1994).

Uncertainty factors for components, say, u_1, u_2,…, u_n, of a product combine to yield the uncertainty factor for the product via the formula:

Uncertainty factor for product
$= \exp\{2*[(0.5*\ln(u_1))^2 + (0.5*\ln(u_2))^2 + \ldots + (0.5*\ln(u_n))^2]^{0.5}\}$

(based on approximating the normal distribution on the log scale as a sum of normal distributions for the different components, each with an approximate 95% probability interval of 2 standard deviations.) For example, the uncertainty factors of 2, 5, and 10 Table 3 combine according to this formula to give a total uncertainty factor of 18 for their product. This uncertainty factor approach is used only to make uncertainty calculations more transparent. Monte Carlo uncertainty analysis is more flexible and general, but is less easy to verify by manual checking of calculations.

Technical Note: Monte Carlo uncertainty analysis. Instead of uncertainty factors, conditional probability distributions for model variables can be calculated via exact algorithms in the *Bayesian Network* (BN) formalism discussed in Chapter 3 (Zhang, 1998; Dechter, 1999). A simpler approximate method, widely applied in health risk assessment, is Monte Carlo simulation (Cheng and Drudzdel, 2000). If no Bayesian inference is required, then computer-aided risk assessment tools such as @RISK™, Analytica™, and Crystal Ball™ can be used to randomly sample values from the probability distributions of input variables and to propagate them forward through the deterministic formulas and conditional probability look-up tables (CPTs) for the other variables in a risk model to create approximate distributions for the values of all other variables. If Bayesian inference is to be used to condition on data while propagating input distributions to obtain output distributions, then specialized software such as the Bayesian Net Toolbox or WinBUGS can be used to perform the more computationally intensive stochastic sampling algorithms (typically, Gibbs Sampling and other Markov Chain Monte Carlo (MCMC) methods) required for accurate approximate inference in BN models (Cheng and Drudzdel, 2000; Chang and Tien, 2002; Andrieu *et al.*, 2003).

Table 3 has illustrated the calculations and an intermediate result for one risk management decision option: doing nothing, i.e., the *status quo* option. In other words, 1.84 cases per year is a preliminary upper-bound point estimate (ignoring uncertainty factors, health consequences, and preventable fractions, as explained above) of the number of severe macrolide-resistant *C. jejuni* cases that may be deprived of clinical benefits of macrolide treatment each year in the absence of intervention. However, as previously emphasized, to support rational decision-making, *it is essential*

to evaluate more than one option, i.e., to inform the decision-maker about the consequences of alternative choices. Policy makers do not always heed this principle. For example, one common alternative approach to decision-making is to apply "situation-action" rules in which surveillance and monitoring data can trigger pre-specified interventions whenever certain conditions are detected, without comparing the probable human health consequences of alternatives. However, risk analysis generally strives to support "rational" decision-making, i.e., decisions made by comparing the likely consequences of alternative decision options and choosing the one that yields the most preferred achievable probability distribution of consequences. (See Cox, 2001, Chapters 5-7 for a survey of rational choice theory and decision analysis for risk analysts.) This book focuses on the use of risk analysis to support such rational risk management decision-making.

The simplest alternative to the "do nothing" (i.e., *status quo*) option is to restrict or ban macrolide uses in chickens. The probable human health consequences of such an intervention, measured by the incremental number of illness-days caused or prevented, show the human health risks or benefits, respectively, of this option. As explained in Chapter 8, a similar framework to that in Table 3 can be used to calculate the human health benefits of animal antibiotic use for the *status quo*, i.e., the human health harm per year now prevented by continued use of animal antibiotics that would occur if such use were terminated. In Chapter 8, continued use of macrolides is estimated to potentially prevent thousands of *C. jejuni* cases per year by promoting animal (and hence human) health and increasing microbial safety of meats. This potential reduction in human health risk significantly outweighs the 1.84 potential severe macrolide-resistant cases per year estimated in Table 3 that might be preventable by changing the *status quo*.

The RRRT example calculations in Table 3 illustrate basic calculation methods and data that can be used for antimicrobial risk assessment. More sophisticated techniques have been developed by and for specialists. For example, quantitative risk assessments often apply results from probability and statistics to calculate the probabilities of conjunctions of events (e.g., as products of marginal and conditional probabilities, extensively used in the RRRT approach; see Chapters 6 and 8). In addition, applied probability and statistics provides many "limit laws" and accompanying statistical procedures that allow risk assessors to estimate probability distributions of population risks (to a close approximation in large populations) based on very partial knowledge of the probability distributions of the factors that contribute to them. For example, as covered in many text books on applied probability and statistics (e.g., Feller, 1968; Lange, 2003):

- *Rare events,* including sporadic cases of foodborne illness, often obey a Poisson approximation law (e.g., Barbour, 2000). Statistical methods

such as Poisson regression can then be used to estimate conditional event rates from data and to test whether a Poisson model is appropriate.

- *Sums and averages* of independent or almost independent variables (e.g., total population risks or average individual risks) typically approach normal distributions in large populations, under the conditions of any of several Central Limit Theorems.
- *Products and networks* of calculations often give results with approximate log-normal distributions (Druzdzel, 1994).
- *Extreme values* (e.g., maximum or minimum values) in large populations or in large numbers of independent trials in many situations follow special parametric asymptotic distributions (e.g., extreme value or Gumbel distributions). Similarly, runs of large or small values and times between successive record values also follow special distributions.
- *Deviations* around expected values often have the property that "large deviations" are much (exponentially) rarer than smaller ones.
- *First-passage times* or first occurrence times in probabilistic transition networks often obey "sharp-transition" laws and 0-1 laws.

Such results can allow population risks to be approximated with useful accuracy for large populations and complex models even when there is considerable uncertainty about the values or probability distributions of individual factors in the models. Although this book emphasizes the application of basic calculation methods and data-driven empirical estimates, limit laws and sophisticated statistical estimation methods based on them allow risk assessment methods to be applied to many common practical situations where quantifying approximate probability distributions for the health consequences of alternative actions suffices to improve decisions.

6. RISK MANAGEMENT

6.1 Definition of Risk Management

Formal risk management is a decision process that maps available risk assessment information about the probable consequences of acts, along with value judgment and priority information, into choices of which acts to take. Acts available to risk management decision-makers and policy makers usually include collecting additional information to reduce uncertainty about exposures and risks, as well as opportunities to disseminate existing information and warnings and to require or constrain individual activities.

Risk management decision models can be used to quantify the expected value of additional information (VoI) for improving decision-

making (see Chapter 5), and hence can help to set research priorities. They also prescribe interim decisions to be made unless and until additional information becomes available.

6.2 Purposes and Outputs of Risk Management

Risk management decision processes and institutions are used to prevent, mitigate, transfer, share, and spread risk and to assign liability and compensate victims of risks. Societal options for risk management often include some of the following approaches:

- *Warn:* Inform or warn potential participants about risks of activities and transactions. For example, putting warning labels on food products (e.g., stating minimum cooking requirements) may help consumers take care to avoid risky storage, preparation or consumption practices. Certifying that a product, process, or facility meets specified microbial quality standards can help provide information to the market, allowing some differentiation of food products and informed consumer choice.
- *Facilitate* voluntary risk management agreements (e.g., between producers and consumers, importers and exporters, etc.) by verifying and publicizing relevant risk information, e.g., by defining standards and grades for food products.
- *Insure:* Underwrite producer costs of complying with new standards (e.g., from inspecting and rejecting food products based on new microbial quality standards). Consumer health insurance may also affect the quality of medical care sought and provided for foodborne illnesses.
- *Regulate:* Restrict voluntary activities or transactions (e.g., production, sale, or use of antimicrobial feed additives or other animal antibiotic uses) by imposing constraints, standards, and regulatory requirements based on risk information.
- *Litigation and process design:* Design and enforce processes and rules to help bolster the microbial safety of foods (e.g., tort liability rules; inspection, labeling or licensing programs; worker compensation)
- *Compensate:* Compensate known or suspected victims of hazardous activities, or compel others (e.g., their known or suspected injurers, or tax payers) to pay compensation.

6.3 Methods for Risk Management Decision-Making

Formal methods for risk management decision-making apply the methods and frameworks of decision analysis, optimization, and group decision-making to clarify value trade-offs among competing goals and to select risk management options that correspond to preferred probability

distributions of consequences. This can be done if relative preference weights, called utilities, can be assigned to different possible consequences of the risk management decisions being evaluated (e.g., to different values of the individual and population risk metrics). The probabilities of these different consequences for different risk management decisions are obtained from the output of the risk assessment. Each risk management decision being considered leads to a corresponding set of probabilities for different consequences and their utilities. The decision leading to the greatest mean value of the utility is recommended.

In practice, this formal decision-analytic approach is seldom directly applied. Different participants may have different preferences for outcomes, be willing to make different trade-offs among goals (e.g., minimizing average risk *vs.* reducing inequities in the distribution of risks), and have different tolerances for accepting risks. In such cases, agreed-to utilities for different consequences may not exist, and risk management decision-making requires negotiation and compromise as well as analysis and deliberation. But even when the formal process of decision analysis is not directly applicable, its conceptual framework is still useful for organizing analysis and deliberation, separating beliefs from preferences for consequences, and identifying and resolving relevant conflicts and/or uncertainties about facts and values. Byrd and Cothern, 2000 and Cox, 2001 further discuss individual and group decision-making processes and frameworks for risk management decision-making.

6.4 Methods of Risk Management to Avoid

Well-informed and effective risk management decision-making, i.e., risk management that is likely to produce desired consequences, requires considering *all* of the most important impacts – good and bad – that an intervention is likely to create. Unfortunately, to date, many antimicrobial risk assessments have ignored the human health risks that proposed risk management interventions might *create,* focusing instead entirely on the human health risks that they might reduce or prevent. This represents a breakdown in sound risk assessment and risk management, similar to assessing financial risks of an investment or acquisition based on only one side of a balance sheet. In general, rational risk management requires considering and comparing the *total* human health consequences, both favorable and adverse, of the risk management decision options being evaluated. Risk characterization should provide risk managers with a balanced accounting of the illnesses or adverse human health effects (and other adverse consequences of interest for decision-making) that a risk management intervention might *cause*, as well as of those that it might

prevent. As illustrated in Chapters 6 and 8, the same basic format and logic (multiplicative modeling) can be used to do both.

Risk management decision processes that recommend interventions based on the *status quo* and/or based on beliefs about what might constitute "precautionary" risk management should be avoided if they do not explicitly identify and compare probable human health consequences of alternative decision options. They violate important normative principles of rational and effective decision-making designed to bring about desired consequences. A more effective approach, according to widely accepted principles of decision analysis (see e.g., Cox, 2001, Chapters 5-7 for a review for risk analysts) is to use quantitative risk assessment information about the *probable consequences* of alternative interventions to eliminate dominated options and to choose the best among those that remain.

6.5 Validating Risk Management Results

A risk assessment model predicts the probable human health effects and other consequences of different risk management actions by predicting their impacts on human exposures to hazards. To maximize learning, these predictions should be tested both before and after implementation of a risk management decision. This is done by conducting an *evaluation study* to assess whether the predicted changes in exposures and health effects actually occurred. If not, the risk assessment model may need to be refined (see Validation of Risk Characterization Results in Section 4 of Chapter 5) and the recommended risk management decision may have to be revised.

7. COMMUNICATING RISK ANALYSIS RESULTS

Risk communication facilitates the effective participation and interaction of technical experts, stakeholders, and decision-makers in risk management decision processes and deliberations. Risk communication is also used to present the results of risk analyses to stakeholders, decision-makers, participants, and other audiences. Communication and deliberation drive much of the risk management decision process in many cultures and are essential for successful outcomes. Web resources are available on the sub-field of risk communication within risk analysis generally and for food safety risk communication in particular. Examples include the following:

- http://www.belleonline.com/oct_02.pdf
- http://www.foodsafetynetwork.ca/risk.htm#communication
- http://www.sirc.org/publik/revised_guidelines.shtml

7.1 Steps in Risk Communication

Risk communication should usually begin early in the processes of risk assessment and risk management. A well-planned, thorough approach to communicating risk analysis results typically includes the following steps:

(a) Identify explicit goals and define explicit (preferably, measurable) criteria for success of the communication effort. The goals specify the purposes (the *why*) of the communication effort, while the criteria specify how success is to be defined and assessed.
(b) Identify target audiences for the communication effort (the *whom*).
(c) Select messages to be shared (presented, discussed, etc.) with each audience to achieve the goals. These messages address *what* will be communicated, e.g., "take-home" messages that should be retained by the audience.
(d) Select framing, presentation media, displays, exhibits, interaction styles and formats, and scripts for presenting key messages. These address the *how* of risk communication.
(e) Implement the risk communication plan; and
(f) Monitor results and incorporate feedback about the effects of the communication into a revised plan.

Omitting any of these steps can compromise the effectiveness of risk communication, no matter how strong the rest of the risk analysis.

7.2 Purposes of Risk Communication

The most common goals for risk communication programs are: to *inform* individuals and groups about risks so that they can make better-informed decisions or seek more information; to *influence* people to change their behaviors, their attitudes and beliefs about hazards, and their acceptance of risk management decisions and policy recommendations; to *involve* affected parties in the decision process; and to *facilitate* their participation in conflict-resolution, consensus-building, and collective decision-making about risk management. The field of risk communication provides guidelines, derived mainly from experience, analysis of survey data, and experiments, for how to accomplish these goals by sharing risk information among stakeholders and decision-makers.

A group convened by OIE (Vose *et al.*, 2001), in a general description that could be adapted to apply to many business, organizational, and political discussions and processes, asserts that "The goals of risk communication are the following:

- to promote awareness and understanding of the specific issues under consideration during the risk analysis process, by all participants
- to promote consistency and transparency in arriving at and implementing risk management decisions
- to provide a sound basis for understanding the risk management decisions proposed or implemented
- to improve the overall effectiveness and efficiency of the risk analysis process
- to strengthen working relationships and mutual respect among all participants
- to promote the appropriate involvement of all stakeholders in the risk communication process
- to exchange information on the knowledge, attitudes, values, practices and perceptions of stakeholders concerning the risks in question."

However, in public health decision-making, there is often a tension in risk communication efforts between *informing* and *influencing* or manipulating target audiences (Ng and Hamby, 1997). Risk communication programs are often designed and evaluated based on their success in changing individual behaviors, e.g., by persuading people to stop eating fish from polluted lakes, to start using sun block, to participate in vaccination programs, to wear seatbelts, or to refrain from smoking. Other risk presentations have as their main goals to make decisions that have already been reached palatable to those affected (often a lost cause if those affected did not participate in the decision or do not perceive the decision process as fair and legitimate) and to confer legitimacy on decision processes by holding open meetings and sharing information.

Effective communication and facilitation about food-related risks enables stakeholders, experts, and decision-makers to participate more effectively in risk management decision processes. It does so by structuring how their beliefs, values, and concerns are elicited, shared, used to create and evaluate decision options, and acted on. It may also enable the facilitator to pursue policy goals by setting the agenda and managing the process to promote certain ends.

7.3 Desired Outputs of Risk Communication

A successful risk communication program summarizes and presents the results of risk analysis in a way that clearly and credibly answers the following questions for the intended audience: (a) What should I do now? (b) Why?/What are the benefits? (d) Why should I believe this?

The output of a risk communication program should be an exposition of risk analysis results that is both accurate and effective in changing (or informing) beliefs, attitudes, and/or behaviors. Communication and presentation styles that are most effective in changing behaviors typically differ in structure, content, and emphasis from those that best express the technical content of risk assessment findings or those that best invite and elicit fruitful participation and interaction. For example, accurate communication of technical findings about risks and uncertainties to technically trained decision makers, and effective internal communication about facts, assumptions, conclusions, and uncertainties among expert members of a risk-assessment or risk management team, can greatly benefit from technical methods. Bayesian Network models, simulation-based what-if analyses, sensitivity analyses, risk profiles, and Bayesian posterior distributions can convey precisely what is known, how it is known, and what remains unknown or assumed… to audiences well trained in such methods.

But technically accurate risk communication does *not* necessarily address other key goals effectively, such as telling people what has been decided, or what they should do, in a way that is likely to win agreement and change behaviors (Blaine and Powell, 2001). It may not even give non-specialists the information they need to make improved decisions. It does not address the need to elicit stakeholder concerns and values or to address them in risk assessment and decision making. By contrast, persuasive communication about risks and risk management decisions to stakeholders, media, and the public requires different skills and emphases, including: building trust, gaining and maintaining credibility and perceived legitimacy, and preparing effective summaries of decision-relevant information using appropriate framing techniques. Brevity, clarity, focus, candor, cogent examples, and deliberate attempts to distance one's self from negative stereotypes of risk communicators may be crucial for communicating technical risks to non-specialist audiences so that the message is listened to instead of being tuned out or dismissed (e.g., Byrd and Cothern, 2000, Chapter 12.) These factors help to establish an audience's perception of knowledge and expertise, openness and honesty, and concern and care – all of which, in turn, tend to promote trust in the speaker and acceptance of his or her risk messages. More generally, audience members consider the source of information, emotional style, framing, and imputed motives of the speaker in assessing the credibility of the message and in responding to it along the continuum from outrage to acceptance (Sandman, 1993; Chartier and Gabler, 2001).

7.4 Methods for Risk Communication

How risk information is formatted and presented can greatly affect how recipients assimilate and act on it. For example, in medical decisions, people are more likely to elect a medical procedure when it is described as "99% safe" then when it is described as having "1% chance of complications" (Gurm and Litaker, 2000). Presenting relative risks rather than absolute risks and using loss framing instead of gain framing make it more likely that patients will adopt screening procedures. In presenting chemical risks, the language used to describe risks may trigger speculations about the presenter's motives and undermine his or her credibility with the target audience (MacGregor *et al.*, 1999). Understanding such effects can help in preparing the presentation of factual information in ways that are likely to elicit desired responses.

A striking insight from the framing literature is that *there may be no neutral way to present risk information.* Any presentation carries with it potential presentation and framing effects and biases that may affect the recipients' attention, interpretation, and actions. Presenting the same information in different ways and emphasizing fact-rich displays (e.g., cumulative risk profiles) that are not strongly associated with known presentation biases may come as close as possible to providing the information needed for rational decision-making without influencing the decision. Such displays often lack the brevity and focus that are most effective in action-oriented presentations.

Effective risk communication must be concerned with process as well as with outcome. If people believe that identifiable groups are having risks imposed on them unfairly by identified others having superior power, authority, or information, the result is likely to be outrage (Ng and Hamby, 1997). Unresolved outrage can quickly destroy the chances for joint problem-solving as an approach to risk management decision-making and conflict resolution. To resolve such situations, it is important to acknowledge and address the perceived unfair situation, either by correcting it or by discussing how decisions *should* be made when values and interests genuinely conflict and then demonstrating willingness to abide by agreed-to principles of fairness in deciding and communicating what will be done.

The following guidelines for communicating regulatory risk analyses and risk management decisions to the public are representative of much prescriptive literature on structuring risk communication and management efforts (e.g., Ng and Hamby, 1997).

Elements of a Successful Agency Risk Communication Plan

1. *Be clear on the roles and goals* of the risk management program (e.g., is the goal to inform, influence, or involve the audience?)
2. *Address stakeholder concerns.* What knowledge, beliefs, values, attitudes, cultures, and contextual factors shape their concerns and motivate their actions?
3. *Study/understand risk perceptions*, concerns, and most effective communication styles.
4. *Involve stakeholders.* Successful risk communication should be interactive and participatory, not a one-way broadcast.
5. Develop technical risk assessment content to support effective risk communication by *answering specific questions*/addressing specific concerns. Emphasize decisions and consequences, not pure science.
6. *Organize risk assessments to facilitate effective presentation* of content. Identify outcomes of interest or concern to stakeholders, identify decision options, show how they affect outcome probabilities, and quantify trade-offs among likely consequences of different options.
7. *Organize risk management decision processes* to eliminate outrage, accomplish goals, serve chosen roles, and reflect Agency values.

7.5 Participatory Risk-Management Decision Processes

Risk management and risk communication intersect in societal risk management decision processes. Public health and regulatory risk management decisions affecting food safety are usually made by multiple participants and reflect the interests of multiple stakeholders with partially conflicting interests and beliefs. The participants interact through *decision processes* in which individual proposals, choices, offers, commitments, and actions or behaviors are iteratively modified until an outcome is reached. In general, risk management decision processes refer to procedures by which multiple participants jointly determine how risks are to be managed. Each participant uses information about what others have done, claimed, or offered to decide what to do next. Their interacting decisions determine how risks are managed.

Aspects of a risk management decision process that often contribute to its perceived legitimacy, and hence its potential effectiveness in changing people's attitudes and behaviours, include how well it does each of the following:

- *Identify and involve key players* (or "stakeholders") early on whose expertise, participation, assent or consent will later be needed.

- *Give each stakeholder opportunities and a positive incentive* to participate (e.g., an expectation of helping to make collective choices).
- *Allow individual concerns, preferences and values to be surfaced, acknowledged, and responded to.* Confront and resolve conflicts among individual beliefs and/or preferences using stated principles for how decisions should be made when individuals disagree.
- *Partner with stakeholders to build trust* in the process, get it used, and improve it over time.

Techniques for managing group dynamics and for organizing and running meetings and hearings effectively (e.g. Bazan, 1998) can often create a broadly shared perception that most of these elements have been accomplished. For example, giving all stakeholders a visible, public opportunity to comment; recording and systematically responding to (or at least noting) points raised; and actively encouraging participation are simple methods that go far toward making a process look and feel legitimate, even if the information collected is subsequently disregarded or poorly used, and even if the final decision reached is unlikely to produce desired consequences. Allowing participants to take turns speaking, keeping and publishing notes and written responses to questions and issues raised, and providing multiple opportunities to review and comment before a final decision is made can help to create perceived legitimacy for public risk management processes.

The importance of participatory risk management and communication for effective risk management decision-making is now well recognized and broadly accepted in many frameworks for risk analysis. For example, based on "a comprehensive analytical review of the risk assessment, risk management, and risk communication approaches currently being undertaken by key national, provincial/state, territorial, and international agencies", Jardine *et al.* (2003) offer "the following 'checklist' to ensure that a good risk management decision is proposed:

- Make sure you're solving the right problem.
- Consider the problem and the risk within the full context of the situation, using a broad perspective.
- Acknowledge, incorporate, and balance the multiple dimensions of risk.
- Ensure the highest degree of reliability for all components of the risk management process.
- Involve interested and affected parties from the outset of the process.
- Commit to honest and open communication between all parties.
- Employ continuous evaluation throughout the process (formative, process, and outcome evaluation), and be prepared to change the decision if new information becomes available."

All group decision processes for risk management have some intrinsic limitations. For example, those who set the agendas for group decision processes and process the results may be able to manipulate the probable outcomes even for decision processes (e.g., voting) that are widely perceived as fair and legitimate. If there is private information, then strategic misrepresentation of interests and beliefs may also hamper the success of decision processes in obtaining fair, efficient outcomes with high probability. Approaching risk management decision processes as exercises in *joint problem-solving* by the participants, backed by a commitment to use mutually agreed-on principles and procedures (e.g., of fairness or voting) to resolve conflicts when necessary, provides a powerful practical approach for creating consensus and acceptance of outcomes despite these potential limitations.

8. SUMMARY AND CONCLUSIONS

This chapter has suggested several goals and criteria for successful risk analysis, meaning risk analysis that can help to promote improved risk management decision-making. A successful risk analysis should do all of the following:

- *Scope the analysis to support decisions* by estimating the causal relations between decisions and probable resulting exposures, and between exposures and their probable total human health consequences. To guide rational regulatory decision-making, traditional quantitative risk analysis seeks to quantify the uncertain *causal relation* between regulatory actions that might be taken and their total probable human health consequences.
- *Evaluate proposed solutions, not problems.* The risk analysis should yield evaluations and comparisons of proposed risk management *actions and interventions*, not simply descriptions of the current situation. A successful risk analysis shows the estimated changes in frequencies and magnitudes (and uncertainties) of human heath consequences resulting from different proposed risk management decisions. It is important to identify an adequate range of risk management options to assure that dominant alternatives are not overlooked.
- *Evaluate total human health impacts.* Total health consequences are found by summing the impacts of proposed actions on human exposures to microbial loads of bacterial species (both resistant and susceptible) over all relevant pathways that contribute significantly to the outcome (e.g., different food animal species, drinking water, home-cooked meals,

restaurant dining, etc.) Applying an exposure-response model to the changed exposures for different decisions then yields the estimated risks associated with them (see Chapter 5).

- *Communicate clearly and enable effective participation.* A well-conducted risk analysis enables its recipients to participate more effectively in risk management deliberations and to communicate questions and concerns more clearly and concisely than would otherwise be possible. It does so by providing them with the relevant information needed to determine the probable consequences of proposed actions and by showing how sensitive these predicted consequences are to specific uncertainties and assumptions in the analysis.

Bailar and Travers (2002) offer additional pragmatic criteria for a useful antimicrobial risk assessment approach. They suggest that such an approach should promote: reduced demand on resources; a common format; reduced demands for data; easy comprehension by non-experts; and ready adaptation. In pursuit of these goals, they recommend a multiplicative model, similar in concept to the RRRT framework illustrated in Table 3. They state that such a model should estimate the annual number of symptomatic infections by the organism of interest in a specific population; the fraction of those occurrences in which the bacterial strain was clinically resistant to the antimicrobial or class of antimicrobials under study; the annual number of occurrences in which infection by a resistant strain led to the specific adverse health outcome(s) under study; and the fraction of the adverse outcomes in which the antimicrobial resistance was a result of the farm use or category of uses under study. In addition, as discussed in the following chapters, it is important to consider the fraction of outcomes that would be prevented (or caused) by a change in animal antibiotic use.

This page intentionally blank

Chapter 3

Hazard Identification

This chapter expands upon the risk assessment portion of the risk analysis process. While Chapter 2 focused on the desired *outputs* of risk assessment, this chapter and Chapters 4 and 5 emphasizes the *process and methods* for producing those outputs. After reviewing the definition of risk assessment and introducing a Bayesian Network formulation of the risk assessment process, we discuss and illustrate the traditional steps of the process: hazard identification, exposure assessment, dose-response modeling (sometimes called hazard characterization), and risk characterization (including uncertainty analysis and sensitivity analysis.) Of these steps, hazard identification, which requires drawing conclusions about causation from data, often presents the greatest conceptual and technical challenges for both microbial risk assessment and analysis of the human health effects of animal antibiotic uses. It is the focus of most of this chapter. Chapter 4 addresses exposure assessment and Chapter 5 discusses dose-response modeling and risk characterization

1. DEFINITION OF RISK ASSESSMENT

The *Codex Alimentarius* Commission, an international (FAO/WHO) body which compiles international food codes and related information, defines *risk assessment* as "A scientifically based process consisting of the following steps: (i) Hazard identification, (ii) Hazard characterization, (iii) Exposure assessment, and (iv) Risk characterization." This chapter defines, explains and discusses the first of these steps. Throughout the discussion, "scientifically based" is interpreted to mean: "Based on specifically identified, independently verifiable data sources and on explicitly stated,

empirically testable hypotheses, models, and calculation formulas or algorithms." Information on the *Codex Alimentarius* Commission and its work related to food safety and risk assessment, including microbial and antibiotic risk assessments, can be found on-line, starting from these sites:

- www.codexalimentarius.net/web/index_en.jsp
- www.fao.org/DOCREP/005/Y2200E/y2200e07.htm

As discussed in Chapter 2, risk assessment of animal antibiotics is used to predict the probable human health consequences of making changes in animal antibiotic use. It predicts the probable change in the number of illnesses or other adverse events per year (or per capita-year, for individual risk) that will be caused by a change in animal antibiotic use. Hazard identification deals with how to establish cause-and-effect relations from data. Exposure assessment quantifies the changes in exposures caused by changes in animal drug use, while hazard characterization (better known in other areas of risk assessment as dose-response modeling or exposure-response modeling) quantifies the causal relation between changes in exposures and probable resulting changes in adverse consequences. Finally, risk characterization integrates the preceding components to predict the probable changes in health that will be caused by a risk management action that changes exposures.

2. A BAYESIAN NETWORK FRAMEWORK

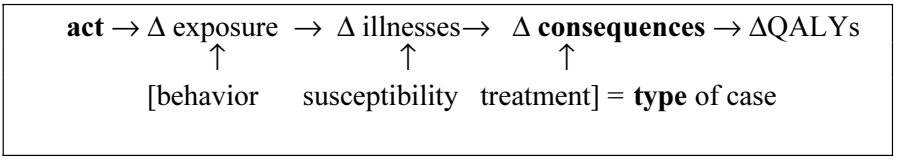
To support effective risk management decisions, human health risk assessments must characterize known or suspected potential causal relations between risk management actions (including the *status quo* or "do nothing" option) and probable resulting human health consequences. The actions typically affect exposures to sources of risk (i.e., hazards), while the consequences typically include changes in frequency or severity of resulting illnesses or deaths in affected populations.

Impacts of changing animal antibiotic uses can potentially be transmitted to humans by several causal paths, such as resulting changes in exposures to both susceptible and resistant strains of bacteria in food commodities, and perhaps transfer of resistance determinants to humans via these or other bacteria. The medical consequences of changes in exposures to microbial hazards depend on resulting changes in illness rates, on patterns of resistance to human drugs among cases of food-borne illness, and on treatment and prescription patterns for patients receiving human antibiotics. Hazard identification identifies causal relations (possibly including identifying causal paths) leading from risk management actions to their human health consequences. Hazard identification often precedes any plan

to develop a risk management strategy, as effective risk management is often impossible if causal relations are not present or are not understood.

Figure 1 outlines a causal graph (Shipley, 2000; Greenland and Brumback, 2002) for assessing risks to humans from changes in animal drug uses. In this template, risk management *actions* that change current practices or activities such as animal drug use can thereby change *exposures* of individuals to potentially harmful agents (the hazards, typically one or more bacterial strains). Changes in exposures, in turn, change expected *illness rates* and hence adverse *health consequences* (e.g., illness-days or early deaths per capita-year) in susceptible members of the exposed population. If desired, different human health consequences can be aggregated into a single summary measure, such as quality-adjusted life-years (QALYs) (Hazen, 2003), but this is optional. Effects of changes in animal drug use on QALYs lost per year in the population may be mediated by behaviors (e.g., kitchen hygiene, cooking, and care-seeking behaviors), individual attributes (e.g., immune status, age, sex, and other covariates that affect susceptibility to infections), physician prescription behaviour, and hospital infection control practices. These covariates may also influence each other (indicated by the brackets [] around them in Figure 1.) For example, an AIDS patient may have food consumption and preparation behaviors and receive medical care and prescriptions that differ from those of a non-AIDS patient. Risk assessment helps to identify risk management options (acts) that decrease adverse health consequences, taking into account the distribution of covariates in the population.

Figure 1: A Causal Graph for Health Risk Analysis



Technical Note: Bayesian Network risk model. A useful mathematical and statistical framework for Figure 1 interprets it as a Bayesian belief network (BN) or causal graph model (Greenland and Brumback, 2002;Chang and Tian, 2002). Each variable with inward-pointing arrows is interpreted as a random variable with a conditional probability distribution that depends only on the values of the variables that point into it. Because this diagram has a decision node ("act") and a value node ("ΔQALYs"), it is an example of an influence diagram (Owens *et al.*, 1997). The essence of the forward Monte Carlo approach to modeling and evaluating uncertain risks in this framework is to sample successively from the conditional distribution of each variable, given the sampled values of its predecessors. Important micro-biological processes, such as cross-contamination during processing, are represented

only implicitly, e.g., by conditional probability distributions of microbial loads on outgoing (processed) carcasses, given the microbial loads on incoming carcasses. Algorithms to identify possible causal graph structures from data (and hence to test whether hypothesized causal theories are consistent with data) have been developed (e.g., Tsamardinos *et al.*, 2003) but are not yet routinely applied in risk assessment.

Each choice of a risk management **act** in Figure 1 generates an approximately Poisson-distributed (Feller, 1968) random number of incremental illness cases ("responses") caused or prevented each year in each severity class (e.g., mild, moderate, severe, fatal) in the population (and in each subpopulation, if there are several). The expected health consequences of this change can be calculated from the following three components models, which are common to most risk assessments:

- An *exposure model* (the "act → Δexposure" link in Figure 1). For microbial risk assessment, this link quantifies the average number of contaminated units (e.g., servings or bacteria, depending on the dose-response function, as discussed in Chapter 2) ingested per year, for population risk; or average contaminated units ingested per capita-year, for individual risks. "Contaminated" here means carrying enough pathogenic bacteria (possibly just one) to pose an elevated risk of food-borne illness to susceptible consumers. The number of contaminated units ingested per year is typically Poisson-distributed, and so is fully characterized by its mean. The exposure model may depend on a consumer's "type", i.e., on individual covariates such as food purchasing, preparation and consumption variables that affect exposures.
- A *dose-response* or *exposure-response* model (the "Δexposure → Δillnesses" link in Figure 1) that quantifies the probability of illness or expected number of cases of each given severity (for infectious illnesses) per contaminated unit ingested. In general, this relation may also depend on the consumer's "type", i.e., on the combination of covariate values that affect risk for that individual, as well as on the dose ingested.
- A *health consequence model* (the "Δillnesses → Δconsequence" link in Figure 1) quantifying probabilities of different health outcomes (e.g., survival *vs.* fatality, or number of QALYs lost) from each case. These outcome probabilities may depend on physician prescription behavior and hospital infection control standards for different types of cases.

These three sub-models determine the expected illnesses and QALYs lost per year in each severity class, for each choice of act. Multiple bacterial strains, food animals and commodities, and at-risk populations (perhaps including groups receiving different medical treatments) can be included in a risk assessment to quantify the *total* human health impact of changing

animal drug uses. Then, summing health impacts over all distinct combinations of affected bacteria, foods, and target populations (each combination corresponding to an instance of Figure 1) gives the total probable change in human health consequences for the act.

Technical Note: Monte Carlo integration. If there are too many combinations for explicit summation over all of them to be practical, but it is easy to generate random samples from the joint distribution of the variables that determine risk, then Monte Carlo simulation methods can be used to obtain accurate numerical approximations of the average value of risk. For example, suppose that the risk model is: risk = $f(x_1, x_2, \ldots, x_n) = f(\mathbf{x})$, and that it is easy to sample from the terms on the right-hand side of the following identity for the joint probability density function (PDF) of these variables: $Pr(x_1, x_2, \ldots, x_n) = Pr(x_1)Pr(x_2 \mid x_1)\ldots Pr(x_n \mid x_1, \ldots, x_{n-1})$. Then Markov Chain Monte Carlo (MCMC) simulation techniques, including Gibbs sampling (Andrieu *et al.*, 2003; Lange, 2003), can be used to generate random samples from the joint PDF of **x**, i.e., from $Pr(\mathbf{x}) = Pr(x_1, x_2, \ldots, x_n)$. Taking a simple arithmetic average of values of $f(\mathbf{x})$ for a sufficiently large random sample of **x** tuples will give an accurate estimate of the true average risk, $E_{Pr(\mathbf{x})}[f(\mathbf{x})]$ implied by the risk model consisting of $f(\mathbf{x})$ and $Pr(\mathbf{x})$. Commercial risk analysis software tools such as Analytica™, @RISK™, and Crystal Ball™ include forward-sampling Monte Carlo simulation routines (including sampling plans and variance-reduction techniques to reduce CPU time) that can generate estimated means, confidence bands, and entire estimated probability distributions for $f(\mathbf{x})$ and that facilitate specification and documentation of risk models. Vose (1998) and Cassin *et al.* (1998), respectively, provide a basic introduction to Monte Carlo simulation in spreadsheet models for microbial risk assessment and discuss how to use Monte Carlo simulation for tasks such as research priority-setting and risk management decision-making.

The conceptual framework in Figure 1 can be implemented with more or less sophistication. Perhaps the simplest useful approach is to estimate each of the following three quantitative factors by a single number, for each risk management act and path being evaluated:

- *Exposure factor* = contaminated servings ingested per capita-year;
- *Dose-response* factor = expected cases of illness per contaminated serving ingested;
- *Health consequence factor* = expected QALYs lost (or illness-days created, etc.) per case of illness. (Alternatively, a vector of expected numbers of different health outcomes can be estimated, e.g., mild, moderate, severe, and fatal illnesses per case.)

In this "Rapid Risk Rating Technique" (RRRT) approach, each sub-model corresponding to a horizontal link in Figure 1 is, in effect, reduced to a single number. Multiplying these numbers together, and then multiplying by the number of people affected, for each causal path (i.e., for each bacterium-food-human subpopulation combination of interest) for a risk management

action, and then summing the resulting products over all causal paths, provides an estimate of the total human health impact per year for that action. A more refined calculation can be made by considering how factors change over time and then summing over time periods (perhaps with discounting). A simpler expedient is to assess and compare the steady-state equilibrium annual risks for different risk management scenarios after all transients have settled down. Chapters 6 and 8 develop this approach.

At the other end of the sophistication spectrum, instances of Figure 1 can be assessed and applied to risk estimation problems using conditional probability calculation algorithms developed for Bayesian Networks and causal graphs (Chang and Tian, 2002). In this case, *hazard identification* can be thought of as identifying instances of Figure 1 that are consistent with available data. Statistical methods are available to test whether specified causal graph models are indeed consistent with data (Greenland and Brumback, 2002; Shipley, 2000) and practical algorithms have been developed to identify potential causal graph models from multivariate data (Aliferis *et al.*, 2003; Tsamardinos *et al.*, 2003). The remaining steps in the risk assessment process can be interpreted as quantifying and applying the Bayesian Network model. Within the Bayesian Network framework, the simple approach of multiplying exposure, dose-response, and consequence factors together generalizes to allowing arbitrary probability distributions for inputs and conditional probability relations or functions at the nodes to be combined (composed) via Monte Carlo simulation (Andrieu *et al.*, 2003) to derive the joint probability distributions of outputs.

Although the Bayesian Network modeling perspective is potentially very useful in risk assessment, it has not been widely applied in animal antibiotic risk assessment and microbial risk assessment, although this is starting to change (e.g., Parsons *et al.*, 2005). Until recently, many biological scientists have been unaware of statistical methods for specifying and testing causal hypotheses objectively using observational data (Shipley, 2000), leading to frequent reliance on expert judgment as the best known approach. Empirically, however, human judgments are notoriously unreliable in matters of statistical and causal inference and risk management decision-making (Plous, 1993). They often reflect preconceptions (Fugelsang and Thomson 2003), perceived plausibility of envisioned mechanisms (Ahn and Bailenson, 1996; Tangen and Allan, 2004), focus on human error and blame (Morris *et al.*, 1999), and cognitive heuristics and biases (Hagmayer and Waldmann, 2002; Bornstein and Emler, 2001), including evaluating prospects by comparison to inferior (hence, normatively irrelevant) ones rather than based only on their own consequences (Stewart *et al.*, 2003; Schwartz and Chapman, 1999, for medical decisions). These factors can lead to incorrect inferences and predictions, pursuit and

inappropriate use of normatively irrelevant information in decision-making (Bastardi and Shafir, 1998) and ineffective decisions (Elstein, 1999 for individuals; Jones and Roelofsma, 2000 for teams).

Bayesian Network methods and objective statistical tests for potential causality, such as conditional independence tests (Shipley, 2000; Greenland and Brumback, 2002), appear promising for achieving more effective, data-driven risk assessments. They can also be used to encode expert judgments when necessary. Development and use of Bayesian Networks and related methods for risk assessment and comparison to other approaches (such as discrete event simulation modeling) has begun (e.g., Parsons *et al.*, 2005).

3. INTRODUCTION TO HAZARD IDENTIFICATION

Risk assessment begins with *hazard identification*. Hazard identification specifies the scope of the assessment – what specifically will be assessed? – and summarizes empirical evidence that exposure to a specific agent causes specified adverse health effects in exposed individuals or populations. Such agents, called hazards, may include food-borne pathogens that can make consumers sick and resistance determinants that make bacteria resistant to antibiotic treatment, leading to increased illness-days, QALYs lost per case, or other clinical harm.

3.1 Definition of Hazard Identification

The *Codex Alimentarius* Commission defines *hazard identification* for food safety as "The identification of biological, chemical and physical agents capable of causing adverse health effects and which may be present in a particular food or group of foods" (Procedural Manual of the Codex Alimentarius Commission - Twelfth Edition). These agents, i.e., sources of risk, are the hazards. Thus, a *hazard* is defined as a *potential cause of an adverse human health effect.* Potential adverse human health effects or consequences of exposures can include increased frequency, duration, or severity of food-borne illnesses, or treatment failures resulting in clinical harm (e.g., increased duration or severity of illnesses).

Example: Alternative Definitions of Hazard Identification

Problem: A risk assessment team identifies the hazard that it will investigate as: "Antibiotic A-resistant *Campylobacter* in chickens caused by use of antibiotic A as a growth promoter in chickens", where "A" is the name of some specific antibiotic or class of antibiotics (e.g., "A" might be Tylosin Premix or macrolides.) Does this

definition give an adequate description of the hazard to be investigated? Why or why not? Does identifying "human illnesses caused by A-resistant *Campylobacter* with A-resistance resulting from use of A in chickens" as the hazard of concern meet the requirements for a useful hazard identification?

Solution: The phrase "A-resistant *Campylobacter* in chickens caused by use of antibiotic A as a growth promoter in chickens" does not specify any adverse human health effect caused by A-resistance, so it does not adequately define a hazard. Even if the phrase is amended to "human illnesses caused by A-resistant *Campylobacter* with A-resistance resulting from use of A in chickens", it is still not fully satisfactory. For example, suppose, for purposes of conceptual clarity only, that resistance does not affect human health (i.e., resistance has no clinical relevance). Then studying "human illnesses caused by A-resistant *Campylobacter* with A-resistance resulting from use of A in chickens" would *not* be studying a source of human health harm caused by use of A in chickens (since, by hypothesis, A-resistance is irrelevant to human health.). In other words, despite the appeal of the phrase as apparently referring to a hazard, in fact, no hazard has been identified in the sense meant in risk analysis – as a potential cause of an adverse human health effect. A more satisfactory definition of the hazard in this case would be "A-resistant *Campylobacter* in chicken, caused by use of antibiotic A in chickens, that leads to increased loss of QALYs among human Campylobacteriosis patients treated with antibiotic A." In the hypothetical example where it is known that the increase in QALYs lost is zero, the hazard identification would show that there is no hazard and no risk, as common sense requires.

As illustrated in this example, it is important to avoid confusing *descriptions* or *naming* of hazards (real or suspected) with *identification* of hazards based on empirical evidence that the specified agents do in fact cause the specified health harm. Describing a *potential* hazard, i.e., one that may or may not cause adverse effects, can help specify the scope of the problem to be addressed by risk assessment, but is not the same as identifying an actual hazard. Unfortunately, many authoritative discussions of hazard identification for microbial risk assessment (MRA) propagate confusion on this point. For example, the *Codex Alimentarius* Commission states that "For microbial agents, the purpose of hazard identification is to identify the microorganisms or the microbial toxins of concern with food." (http://www.foodriskclearinghouse.umd.edu/pversion/Codex_MRA.htm). Here, the implied definition of hazards as *agents that cause adverse effects* is replaced with a more ambiguous definition as agents that are "of concern" (to whom and for what reasons being left unspecified.) The World Health Organization describes hazard identification as "The identification of known or potential health effects associated with a particular agent" (http://www.who.int/foodsafety/micro/riskassessment/en/). This description drops the crucial ingredient of causality altogether.

This book will use the definition of hazards as agents that can *cause* adverse human health effects, not agents that are of concern for other reasons or that are only statistically associated with adverse health effects but that do not cause them.

3.2 Purposes and Goals of Hazard Identification

Hazard identification serves the following main purposes:

1. *Rapidly screen potential hazards* by identifying whether available data support the hypothesis that specific adverse health effects might be caused by specific exposures or activities. Hazard identification can benefit by applying formal methods of causal analysis (e.g., Shipley, 2000) to determine whether hypothesized causal relations linking acts to exposures and to adverse health responses and consequences are consistent with available data.
2. *Identify causal relations between exposures to specific foodborne hazards and specific adverse human health effects.* To support risk management decision-making, it is often helpful to identify exposures or hazards resulting from controllable decisions or behaviors.
3. *Identify risk factors, behaviors, and exposure conditions that increase risks to specific exposed populations* (e.g., the old, the young, the immuno-compromised, etc.)
4. *Present and evaluate empirical evidence for and against the hypothesis that exposures to specific food-related hazards (typically resulting from controllable decisions, e.g., on use of feed additives) cause specific adverse human health effects.* This is somewhat analogous to the US EPA's statement that, for environmental hazards, "The objective of hazard identification is to determine whether the available scientific data describe a causal relationship between an environmental agent and demonstrated injury to human health or the environment" (http://www.bethel.edu/~kisrob/hon301k/readings/risk/RiskEPA/riskepa1.html).

If the hazard is non-zero, what agent should be considered "the cause" of adverse consequences? Is it the bacteria involved, the resistance determinants that they carry, use of animal and/or human antibiotics that select for those determinants, or failure to properly prepare, cook, or handle food to eliminate contamination? In reality, *joint causation* is common, and the best answer depends on the specific risk management decisions that the hazard identification is intended to support. In general, *any* event or condition, such as exposure to a hazard or decreased immunity to the effects of exposure, that hastens the occurrence of an adverse effect (e.g., by stochastically reducing the random time until its occurrence) can be viewed as a contributing "cause" of the effect. Causes of causes are also (indirect)

causes. (For more on the philosophical definition and ambiguities of "causation", see Williamson, 2005.) Thus, "the cause" of an adverse health effect is often not uniquely defined. Nonetheless, as a practical matter, it often suffices for purposes of effective risk management to be able to predict the effects of alternative risk management interventions on the rates at which adverse events of different severities occur. Risk assessment provides methods to do so. Hazard identification helps to identify interventions that might be able to cause desired changes in human health effects, typically by reducing exposures to hazardous agents.

Example: Propagation of Causal Effects *vs.* "But-For" Causation

In the following diagram, a change in animal antibiotic use (ΔAAU) propagates in the direction of the arrows to cause a change in human illnesses per year:

$$\Delta\text{AAU} \rightarrow \Delta\text{ill flocks} \rightarrow \Delta\text{CFUs per serving} \rightarrow \Delta\text{illnesses per year}$$

That is, a change in AAU causes a change in the fraction of flocks that are ill (e.g., airsacculitis-positive or necrotic enteritis positive at the time of slaughter, for a chicken flock), which changes the microbial load (shifting the frequency distribution of colony forming units, CFUs, of pathogenic bacteria) on meat reaching consumers, which in turn affects the number of illnesses per year among people consuming the food commodity. The "cause" of the change in illnesses per year in this diagram is not unique: *any* of the preceding changes can be identified as a contributing cause. (For example, in a legal argument using "but for" causation, it would be true that the change in ill flocks would have had no impact on illnesses per year if processing or cooking had been completely effective in killing all bacteria, so that ΔCFUs per serving would not have changed. Conversely, the owners of these processes could correctly claim that the change in illness rates would not have occurred "but for" the change in animal antibiotic use that changed animal illness rates.) However, if risk analysts understand how to quantitatively predict effects of changes in this model, they can calculate what change in AAU (or other changes at intermediate points) are needed to reduce illnesses per year by a given amount. They can determine what reductions can be achieved (and how) by optimizing over a given set of feasible interventions. Thus, what is required for effective risk management is not a unique definition of the cause of Δillnesses per year, which does not exist, but only an adequate quantitative understanding of how changing controllable inputs will affect the output of concern, illnesses per year. That is the task of risk assessment.

Ideally, a hazard identification for microbial risk assessment should identify the microorganism that causes specific diseases or adverse health effects (e.g., using Koch's postulates), elucidate the infection and disease

process (including the conditions under which infection and illness occur); identify possible transmission routes (e.g., food, water, vectors); and identify covariates (e.g., host immune status, other risk factors) that can interact with or affect the relation between exposure and risk (Haas *et al.*, 1999). A hazard identification for antimicrobial risk assessment should also identify the causal relation between use of antimicrobial drugs in animals and the levels of resistant pathogens in human patients, as well as the causal relation between these levels and the frequency and magnitude of increased mortality risks, morbidity risks, and treatment failure rates.

If objective statistical tests for hazard identification do not identify a causal relation between decisions (or actions), exposures, and human health risks, then this result should be stated, along with discussion of the statistical power of the tests used for the data examined. In this case, risk assessment can still be carried out, but it becomes *contingent* on the assumption that a risk exists. Such a contingent risk analysis can be useful if it shows that risks are small, by providing a plausible upper bound on the true (conjectured but perhaps non-existent) risk. But it may not be useful for accomplishing other risk analysis goals, such as guiding rational choice among expensive risk management alternatives, if the contingent risk estimates are large enough so that intervention may be warranted.

4. METHODS OF HAZARD IDENTIFICATION

4.1 Causal Inference From Epidemiological Data

Table 1 outlines steps for forming and testing causal hypotheses about exposure-response relations using epidemiological data, such as case-control studies of risk factors associated with food-borne illnesses (i.e., comparing exposure histories of ill cases against those of matched well controls), or surveillance data on illness rates in groups known to have different exposures. As more of these steps are completed, the empirical support increases for an inference that there is a causal relation between exposure and risk. Most statistical methods in epidemiological risk analysis focus on steps 1 to 3, i.e., identifying non-random associations and then eliminating potential biases and confounders as likely explanations. These steps can often be carried out using data from observational studies without requiring direct manipulations and experimental verification of predictions. The main method for doing so is sometimes called the *refutationist approach* (Maclure, 1990, 1991). It consists of systematically enumerating possible competing, non-causal explanations for the observed data, and then eliminating each of these alternative potential explanations (if possible) using statistical tests and available data.

Table 1: Steps to Establish a Causal Exposure-Risk Relation

1. *Identify a statistically significant exposure-response association*. Demonstrate that there is a non-random positive statistical association between exposure histories or events and adverse human health consequences in an epidemiological data set. Case-control, prospective cohort, or other cross-sectional or longitudinal epidemiological data may be used for this purpose.
2. *Eliminate confounding* as a possible explanation of the association. Show that it is *not* fully explained by other (non-exposure) causes such as differences in lifestyle factors, age, or exposures to other confounders (Grimes and Shulz, 2002; Feldman, 1998; Greenland and Morgenstern, 2001).
3. *Eliminate biases from sampling, information collection, and modeling choices as possible causes*. Show that the association is *not* explained by biases in who was selected (as study subjects or as controls) or in how information about them was collected and analyzed. (Choi and Noseworthy, 1992; Deeks *et al.*, 2003)
4. *Test and confirm hypothesized causal ordering and conditional independence* relations among observed values of variables. For example, show that the response is *not* conditionally independent of its hypothesized direct causal predecessors (e.g., exposure), but that it *is* conditionally independent of more remote causal predecessors, given its direct predecessors. (Shipley, 2000)
5. *Confirm efficacy of interventions*. Confirm that changes in the levels of direct causal predecessors (e.g., exposures) are followed by the predicted changes in the levels of the variables that they affect. This may often be done from time series observations, even if direct experimental manipulation is impossible, using methods for interrupted time series analysis, intervention analysis, change point detection, and quasi-experiment design and inference (e.g.., Swanson *et al.*, 2001; Green, 1995).
6. *Identify and elucidate causal mechanism(s).* Explain how changes propagate via one or more causal paths to produce effects. A "causal path" is a sequence of steps in which completion of the earlier steps creates conditions that trigger or increase occurrence rates of subsequent steps. Such steps may be identified from experimental data and/or by applying generally accepted laws.

It is traditional in epidemiology to apply a set of criteria or tests (especially ones usually attributed to Sir Austin Bradford Hill, with numerous variations and refinements) to help judge whether an observed statistical association between exposure and adverse health effects is likely to be causal. They include the strength, consistency, specificity (to what extent is the hypothesized cause present in all and only those cases in which the effect occurs?) temporality (does the hypothesized cause precede its hypothesized effect?), biological gradient or dose-response (is presence of more of the hypothesized cause associated with more of the effect?), biological plausibility, and consistency and coherence with other knowledge

of the statistical associations between exposures and the effects that they are hypothesized to cause (e.g., Weed, 2000; Surgeon General's Report, 2004). Unfortunately, these traditional criteria or tests have no necessary relation to causation, since systematic errors in sampling or in modeling may mislead them, as was well recognized by Hill himself (Phillips and Goodman, 2004). For example, an erroneous statistical model (such as the linear no-threshold model $E(y \mid x) = Kx$, if x and y are statistically independent random variables, each uniformly distributed between 0 and 1) can create an apparent strong, consistent statistical association (e.g., a K value statistically "significantly" greater than zero in this mis-specified model) that may seem biologically plausible if prior beliefs incline strongly in this direction (e.g., if x is interpreted as an exposure that is likely to cause excess risk of adverse response y). In a retrospective case-control study where hypothesized causes always precede observed effects (trivially satisfying a non-demanding version of the temporality criterion), one might then be well on the way toward establishing a case for "causality" using traditional criteria, even if there is in fact no causal or statistical relation between exposures and effects and the only source of association is an incorrect model. A variety of other modeling biases (Table 2), as well as omitted variables and residual confounding (Feldman, 1998), can also create associations that satisfy the traditional criteria in the absence of true causation. Conversely, strong causal relations may be present that do not satisfy these criteria.

Example: Unvalidated Modeling Can Create Spurious Associations

The following 2 x 2 table shows the mean days of illness (days of diarrhea) and numbers of observations in a case-control study of campylobacteriosis. The two dimensions are: foreign-travel related *vs.* domestically acquired cases of campylobacteriosis; and ciprofloxacin-resistant *vs.* ciprofloxacin-susceptible cases. While foreign travel is clearly strongly associated with acquisition of ciprofloxacin-resistant cases (29/77 of foreign travel cases are resistant compared to only 41/570 among non-foreign travel cases), and also with longer mean durations of diarrhea, ciprofloxacin resistance is clearly *not* associated with longer mean duration of diarrhea among domestically acquired cases 6.9 days vs. 6.9 days).

	Ciprofloxacin Resistance	
Foreign Travel	**Yes**	**No**
Yes	8.1 days, n = 29	7.6 days, n = 48
No	6.9 days, n = 41	6.9 days, n = 529

Note: Patients not reporting diarrhea, with continued diarrhea, not able to estimate duration of diarrhea, or not responding to the foreign travel question were not included.

Despite this absence of an empirical relation between fluoroquinolone (here, ciprofloxacin) resistance and days of illness among domestic cases, Nelson *et al.*

(2004) used a statistical model with variable-selection and subset-selection conditions [but without appropriate model diagnostics, validation tests, or sensitivity analyses reported (Greenland, 1989)] to conclude that "persons with ciprofloxacin-resistant infection had a longer mean duration of diarrhea than did the persons with ciprofloxacin-susceptible infection (P = 0.01); this effect was independent of foreign travel." (Engberg *et al.* (2004) used a similar modeling methodology to reach similar conclusions for a different data set.) The surprising conclusion that the effect is "independent of foreign travel", and related policy-relevant conclusions about domestic use of fluoroquinolones in chickens threatening their efficacy in human medicine, were apparently based on a misinterpretation of zero *regression coefficients* in the unvalidated statistical model as indicating statistical independence of foreign travel, even though the *data* clearly show such dependence. When model and data disagree this way, the model's conclusions should be rejected and the model should be refined to better describe the data (Greenland, 1989).

In general, researchers and risk assessors should be very careful not to search for and adopt combinations of modeling assumptions that support specific beliefs but that do not fit the data. Analysts and decision-makers alike should start with healthy scepticism toward model-driven conclusions and not necessarily accept them unless and until appropriate model diagnostics and results of model validation tests (Gelman *et al.*, 2005) and sensitivity analyses have been reported (Greenland, 1989) and corrections for model selection biases and multiple testing biases have been made.

Uncorrected *multiple testing bias*, in which statistical modeling generates false positive results at far higher rates than reported significance levels and P-values suggest when researchers (or computer programs) search among many alternative sets of modeling assumptions and report only apparently significantly positive results, has been too common in published epidemiology studies (Ottenbacher, 1998). Automated stepwise variable selection routines in logistic and other multivariate regression models can generate high rates of false-positive associations even in random data and create spurious non-zero regression coefficients for real-world data sets (e.g., Austin and Tu, 2004). Such errors can be avoided in many ways (Cox, 2001, Chapter 3), e.g., by using non-parametric models, appropriately reduced P-values, re-sampling (cross-validation or bootstrap) estimates of true error rates (Romano and Wolf, 2005), and Bayesian Model-Averaging (Wang *et al.*, 2004). Risk assessors should apply such methods to avoid drawing false statistical and causal inferences. (Technically inclined users of risk assessment can also check that the presented conclusions follow robustly from the underlying data, rather than being artifacts of modeling, before accepting them at face value.) This is a key aspect of the refutationist approach for identifying genuinely causal exposure-response relations.

4.2 Refuting Non-Causal Explanations

Many epidemiologists have recognized that, to draw valid causal inferences, it is necessary to refute competing (non-causal) hypothesized explanations for observed exposure-response associations (Maclure, 1990, 1991). Table 2 summarizes common competing explanations (mainly, confounding and/or sampling, information, or modeling biases) and some suggested statistical methods to refute them (Cox, 2001, Chapter 3.)

Table 2: Potential Non-Causal Explanations for Exposure-Response Associations

Potential Non-Causal Explanations	**Methods to Refute Potential Explanations** (see Cox, 2001 for details)
Modeling Biases	
Variable selection bias (includes selection of covariates in model)	Bootstrap variable selection, Bayesian model averaging (BMA), cross-validation for variable selection.
Omitted explanatory variables (including omitted confounders and/or risk factors)	Include potential confounders in an explicit causal graph model; test for unobserved latent variables
Variable coding bias (i.e., how variables are coded may affect apparent risks)	Use automated variable-coding methods (e.g., classification trees, Lemon *et al.*, 2003). Don't code/discretize continuous variables.
Aggregation bias / Simpson's paradox	Test hypothesized relations at multiple levels of aggregation. Include potential confounders in an explicit causal graph model.
Multiple testing/multiple comparisons bias	Use current (step-down) procedures to adjust p-values (Romano and Wolf, 2005)
Choice of exposure and dose metrics; choice of response effect definitions and measures	Use multiple exposure indicators (e.g., concentration and time). (Don't combine.) Define responses as survival functions and/or transition rates among observed health states.
Model form selection bias and uncertainty about the correct model for exposure-response relation and other relations.	Use flexible non-parametric models (e.g., smoothers, wavelets); Bayesian Model-Averaging (BMA). Report model diagnostics and sensitivity analyses of results to model forms (Greenland, 1989).
Missing data values can bias results	Use data augmentation, EM algorithm, MCMC algorithms (Schafer, 1997)
Measurement and misclassification errors in explanatory variables	Use Bayesian measurement error models, data augmentation, EM algorithm, and other missing-data techniques (Schafer, 1997; Ibrahim *et al.*, 2005)
Unmodeled heterogeneity in individual response probabilities/parameters	Latent variable and mixture distribution models, frailty models of inter-individual variability
Biases in interpreting and reporting results	Report results (e.g., full posterior PDFs) *conditioned* on data, models, assumptions, and statistical methods. Show sensitivities.

Sample Selection Biases	
Sample selection (sample does not represent population for which inferences are drawn)	Randomly sample *all* cohort members if possible
Data set selection bias (i.e., selection of a subset of available studies may affect results)	Use meta-analysis to show sensitivity of conclusions to studies. Use causal graph models to integrate diverse data sets
Health status confounding, Hospital admission bias (and referral bias)	If possible, use prospective cohort design. Use population-based cases and controls (Choi and Noseworthy, 1992)
Selective attrition/survival (e.g., if exposure affects attrition rates) Differential follow-up loss	Use a well-specified cohort. "Include non-surviving subjects in the study through proxy interviews" (Choi and Noseworthy, 1992). Compare counter-factual survival curves
Detection/surveillance bias	Match cases to controls (or exposed to unexposed subjects) based on cause of admission.
Membership bias (e.g., lifestyle bias, socioeconomic history)	• In cohort studies, use multiple comparison cohorts. • Hard to control in case-control studies.
Self-selection bias; Response/volunteer bias	• Achieve response rate of at least 80% by repeated efforts. Compare respondents with sample of non-respondents
Information Collection Biases	
Intra-interviewer bias	Blind interviewers to study hypotheses, subject classifications
Inter-interviewer bias	Use same interviewer for study and comparison groups
Questionnaire bias	Mask study goals with dummy questions; avoid leading questions/ leading response options
Diagnostic suspicion bias Exposure suspicion bias	Hard to prevent in case-control studies. In cohort studies, make diagnosis and exposure assessments blind to each other.

Yet, identifying alternative explanations to be refuted can be perceived as unhelpful when it is confined to merely listing logical possibilities, without addressing their plausibility and likely impacts on risk estimates. For example, Savitz *et al.* (1990) state that "Biases that challenge a causal interpretation can always be hypothesized. …It is essential to go beyond enumerating scenarios of bias by clearly distinguishing the improbable from the probable and the important from the unimportant." They argue that those who do not like a causal interpretation of epidemiological data might readily raise speculative hypothetical potential biases and confounders that can not all be refuted with available resources. This strategy could prevent conclusions about causation from being drawn when common sense and sound policy would be better served by accepting that causation is plausible, even if it is not practical to refute all conceivable alternative explanations. On the other hand, accepting a statistical association as causal without rigorously examining and excluding competing hypotheses may make it too easy to launch expensive risk management control actions that would be

effective if the association were causal, but that will not produce the anticipated benefits otherwise.

A partial solution to this dilemma is to focus on those non-causal explanations that appear to be *likely* and *important* (Savitz *et al.*, 1990) i.e., those (if any) that might plausibly explain most or all of the observed exposure-response associations. Appropriate data analysis methods can often reveal which potential biases and confounders are most likely to provide non-causal explanations in specific studies. They can also help to eliminate logically conceivable biases that do not play a large role. Most importantly, they can help to eliminate the most likely and important non-causal explanations when they are incorrect. Evidence that makes non-causal explanations unlikely makes causal explanations more likely, even if the evidence is not definitive.

In summary, the refutationist approach to hazard identification suggests a key necessary criterion that a claim of causation for an observed exposure-response association must satisfy to be well supported: *have competing non-causal explanations been eliminated?* If so, then the hypothesis of causation is supported by the data used to refute them. Otherwise, the unrefuted potential competing explanations undermine the conclusion that the association is causal (Maclure, 1991).

Example: Does Eating Chicken Cause Campylobacteriosis?

It has long been accepted by many food safety regulators and epidemiologists that the primary cause of campylobacteriosis in the United States and other developed countries is consumption of chicken (e.g., FDA, 2001). The underlying logic is simple: chicken carry *Campylobacter*, people eat chicken, therefore people can get campylobacteriosis from eating chicken. Many peer-reviewed articles cite case-control studies showing significant statistical associations between chicken consumption and risk of campylobacteriosis (e.g., Skirrow, 1991). Thus, hazard identification in this case might seem as simple as stating that "obviously" chicken is a source of *Campylobacter* that can infect people. But the refutationist approach demands more. It requires that associations between chicken consumption and risk of campylobacteriosis must not be fully explained away by confounding.

Data from some recent case-control studies have revealed that, perhaps surprisingly, consumption of chicken meals at home (prepared in any of a variety of ways, as well as purchasing, handling, and preparing raw chicken) are highly statistically significantly associated with *reduced* risk of campylobacteriosis, perhaps due to acquired immunity (Cox, 2002). The relative risk of campylobacteriosis among people eating chicken at home compared to other people is as low as 0.6 in some studies (Effler *et al.*, 2001, Table 1). On the other hand, eating chicken (and other meats) in restaurants is associated with significantly

increased risk of campylobacteriosis (e.g., Friedman *et al.*, 2004). Eating undercooked chicken is also associated with restaurant dining (and with consumption of other undercooked meats) and with increased risk of campylobacteriosis.

Given these patterns, an adequate hazard identification supporting the causal conclusion that chicken is indeed a likely source of campylobacteriosis risk must first refute the alternative hypothesis that any positive association between consumption of (undercooked) chicken and risk of campylobacteriosis is fully explained away by confounders such as restaurant dining (and other commercial dining) and kitchen hygiene. In causal graph terms, it is necessary to show that the correct diagram looks something like model A:

Model A: *restaurant dining* $\rightarrow$ *undercooked chicken consumption* $\rightarrow$ *risk*

rather than like model B:

Model B: *undercooked chicken consumption* $\leftarrow$ *restaurant dining* $\rightarrow$ *risk.*

To empirically test causal graph model A *vs.* causal graph model B using data, it suffices to show whether risk is *conditionally independent* of undercooked chicken consumption, given restaurant dining information (as in model B but not model A) or whether risk is conditionally independent of restaurant dining information given information on undercooked chicken consumption (as in model A but not model B). Shipley (2000) and Frey *et al.*, 2003 discuss appropriate statistical tests. Limited tests done to date tend to support model B rather than model A, suggesting that the "obvious" and plausible hypothesis that chicken consumption is the predominant driver of human campylobacteriosis risk may not be correct (Cox, 2002). Indeed, as noted by Canadian investigators, "Independent risk factors for campylobacteriosis (eating raw, rare, or undercooked poultry; consuming raw milk or raw milk products; and eating chicken or turkey in a commercial establishment) account for < 50% of cases in Quebec. Substantial regional and seasonal variations in campylobacteriosis were not correlated with *Campylobacter* in chickens and suggested environmental sources of infection, such as drinking water." (Michaud *et al.*, 2004).

The preceding example illustrates the potential gap between intuitive judgment-based hazard assessment and formal statistical data-driven hazard assessment. A focus on data analysis and refutationist logic has led some investigators to conclude that chicken consumption has not been established *empirically* as a primary source of Campylobacteriosis in the United States and other developed countries, however appealing the hypothesis may be to common sense, and despite much supporting opinion and many literature

citations for it. (The cited articles generally do not rigorously examine or refute competing explanations and interpretations of the data, however; see Phillips *et al.*, 2004) Many regulators and other stakeholders perceive this empirical, refutationist approach as unconvincing at best and as mistaken, misleading, and obstructionist at worst (e.g., Tollefson, 2004). Opinions are sharply divided on whether refutationist criteria should be met (e.g., by carrying out conditional independence tests to discriminate among competing causal models for the data) *before* acting on the assumption that the common-sense interpretation is correct. Yet, from the standpoint of sound risk assessment methodology, the answer is clear: the burden of hazard identification cannot be adequately met – that is, met in such a way that risk assessment results can be relied on to help inform and improve risk management decisions – without presenting empirical evidence of a true cause-and-effect relation between exposure and adverse effects. Even intuitively "obvious" hazard identification hypotheses made without adequate empirical support often eventually prove to be incorrect, and they should be treated with scepticism until empirical evidence of causation has been supplied and plausible competing hypotheses have been ruled out.

Example: Mortality Associated with Foodborne Bacterial Infections

In contrast to conventional clinical wisdom (e.g., Andrews *et al.*, 1998), a study by Helms *et al* (2003) reported a surprisingly strong, significant association (relative mortality ratio = 2.56) between common foodborne illnesses, including campylobacteriosis and salmonellosis, and increased short term and long term mortality. These authors and co-workers have published subsequent studies with similar methods and findings (e.g., Helms *et al.*, 2005). However, their analysis compared ill cases with healthier (general population) controls. This invites a plausible alternative hypothesis: that sick people are more likely to become victims of foodborne infections and are also likely to die sooner than well people, but underlying illness causes both, rather than foodborne infections causing early mortality. In causal graph notation, the following model B competes with model A as a potential explanation for the reported statistical association between foodborne infections and mortality risk:

Model A: *underlying illness* $\rightarrow$ *foodborne infection risk* $\rightarrow$ *mortality risk*

vs.

Model B: *foodborne infection risk* $\leftarrow$ *underlying illness* $\rightarrow$ *mortality risk*

To help refute this alternative explanation (model B), a "comorbidity index" was used to adjust for the fact that cases were much more likely than controls to be have

serious underlying illnesses (such as AIDS, metastatic cancers, and lymphomas or leukemias). However, the comorbidity index adjustment used is far from perfect. For example, in a study of its use in predicting lung cancer survival, it explained only 2.0% of the variation in survival (Tammemagi *et al.*, 2003). It has not been validated for gastrointestinal infections in AIDS and leukemia patients. Thus, while Helms *et al.* relied on an index that may account for very little of the variance in mortality rates, and attributed the remainder to bacterial pathogens, differences between cases and controls in the prevalence of serious underlying illness might provide a more plausible explanation. This potential confounder was not refuted as a plausible explanation of the reported association. Severe underlying illness (such as AIDS) acts as an *incompletely controlled confounder*, predicting both increased mortality rates and increased risk of campylobacteriosis and salmonellosis infections. The comorbidity index does not fully reflect mortality consequences of illness. Hence, *residual confounding* by illness status, even after conditioning on the index, can create a statistical association between infection and mortality, even if the former does not cause the latter. Unless and until confounding by disease status is ruled out as a plausible explanation for the reported association between bacterial infections and subsequently elevated mortality rates, a causal interpretation of the association would not be justified (Feldman, 1998).

As in the previous example, it is natural to question whether public health is truly best served by insisting on a rigorous refutation of competing explanations before accepting a causal interpretation (model A) for the observed statistical association. However, in this case, the required refutation can be accomplished quite easily (if the suggested causal interpretation in model A is correct) by showing that individuals with the *same* underlying disease status (e.g., otherwise healthy individuals), who differ only in whether they have foodborne infections, differ significantly in their subsequent morbidity rates. This test should be straightforward, and there seems little reason not to apply it.

Example: Caveats for Model-Based Epidemiological Associations

One approach to assessing the potential causal relation between chicken consumption and fluoroquinolone resistance is to examine the associations between them in case-control data. For example, Kassenborg *et al.* (2004) reported that "When patients with domestically acquired fluoroquinolone-resistant *Campylobacter* infection were compared with matched healthy control subjects in a multivariate analysis, those infected were 10 times more likely to have eaten chicken or turkey cooked at a commercial establishment (18 [55%] of 33 case patients vs. 7 [21%] of 33 controls; matched OR, 10.0; 95% CI, 1.3-78). … This study provides additional evidence that poultry is an important source of domestically acquired fluoroquinolone-resistant *Campylobacter* infection."

Such a strong association, if it were truly present in the data set, could be very

useful as a starting point for hazard identification. At a minimum, it would convincingly establish that there is a strong association to be explained. But examination of the raw data in this case (from the same case-control study reported on by Cox, 2002; Friedman *et al.*, 2004; and Nelson *et al.*, 2004, discussed in previous examples) shows that the reported association is *not* actually present in the raw data. Rather, it appears to be an artifact of the specific statistical modeling choices and variables selected, and only achieves apparent statistical significance when the *model uncertainty* (Viallefont *et al.*, 2001) created by this selection is improperly ignored.

Reanalysis of the same data (Cox, 2004) reveals that the reported findings are highly sensitive to the subset of risk factors considered, the choice of variable-selection algorithms (e.g., forward vs. backward stepwise variable selection), the selection of a model form (e.g., logistic regression vs. non-parametric alternatives), and treatment of missing data. The claimed confidence interval for the matched OR excludes 1 (no association) only if uncertainties due to these modeling choices are neglected. Slight variations in modeling approach (e.g., using backward *vs.* forward stepwise variable selection *vs.* Bayesian model-averaging) eliminate the claimed finding of a positive association between fluoroquinolone-resistant campylobacteriosis and poultry consumption. (Also, of course, 55% is not usually considered "10 times more likely" than 21%. The reported matched OR of 10 is not an empirical finding, but only a model-based prediction from an unvalidated logistic regression model, for which appropriate model diagnostics (Bagley *et al.*, 2001) have not been presented.) Thus, *modeling biases* have not been refuted as plausible explanations of the claimed association.

Non-parametric techniques such as classification tree analysis (Lemon *et al.*, 2003) can help to avoid parametric model-selection biases. Kassenborg *et al.* state that "In our final multivariate model, we examined the following risk factors: eating chicken or turkey cooked at a commercial establishment, eating in a non–fast food restaurant, using antacids, and eating nonpoultry meat at home. *Using this model*, we found that eating chicken or turkey at a commercial establishment was the only risk factor that remained independently associated with illness" (emphasis added). By contrast, when we examined the same data set using classification tree analysis, we discovered that exposure to ground beef outside the home and consumption of raw milk both also appear to be potentially significant risk factors for fluoroquinolone-resistant campylobacteriosis. Chicken consumption as a whole and chicken consumption in commercial establishments have non-significantly negative associations with fluoroquinolone-resistant campylobacteriosis.

In summary, the apparent dramatic association presented is highly sensitive to specific statistical modeling choices. Different choices, or use of non-parametric methods to avoid having to make such choices, lead to very different conclusions. Thus, the reported significant strong association between poultry consumption and domestically acquired fluoroquinolone-resistant *Campylobacter* infection appears to

be a theoretical implication of the particular unvalidated model selected by the authors that disappears when less restrictive modeling assumptions are made. It is not a robust feature of the raw data. Therefore, model selection bias remains a plausible, unrefuted alternative (non-causal) explanation for the reported association.

4.3 Seeking Positive Evidence for a Causal Relation

In addition to statistical tests for refuting hypotheses that offer non-causal explanations for observed associations, there are several ways to build positive evidence for a true causal relation when one exists. On the biological side, hazard identification can draw on knowledge of infectious diseases, epidemiological data, and clinical microbiology to create testable hypotheses about causal relations among decisions, exposures, and health consequences (Haas *et al.*, 1999). On the statistical side, hazard identification can apply methods of causal analysis to identify the decision-exposure and the exposure-health effects links in the following causal chain:

$$\text{decisions} \rightarrow \text{exposures} \rightarrow \text{health effects} \leftarrow \text{covariates}.$$

Technical Note: Causal graph terminology. As previously noted, in such a diagram, the conditional probability distribution of each quantity depends only on the values of the quantities that point into it, if any. (See Chapter 1 of Shipley, 2000 or Chapter 4 of Cox, 2001.) A causal chain or graph is *identified* by using statistical tests to show that each node (representing a variable, or quantity) is conditionally independent of all of its more remote ancestors, given the values of the variables that directly point into it (its "parents"). It is *quantified* by specifying the conditional probabilities of each variable's possible values, given the values of its parents, if any. Input variables (having no parents) may have (unconditional) probability distributions reflecting uncertainty about their values. Hazard identification deals with the identification of causal links and models. They are then quantified in the exposure assessment and exposure-response steps.

Categories of objective empirical evidence that are often considered in antimicrobial risk assessments include:

- *Spatial associations* between animal antibiotic use and resistance levels in human patients. Associations between animal antibiotic use and human illness rates may also be relevant if the animal antibiotic use affects microbial loads of pathogens reaching human via meats.
- *Temporal associations* between the date(s) of *introduction* of an animal antibiotic and subsequent changes in animal and human resistance rates (after controlling for contemporaneous changes in other factors and potential confounders, e.g., changes in foreign travel).

- *Temporal associations* between the date(s) of *cessation* of an animal antibiotic (e.g., following the European bans on growth promoters) and subsequent changes in animal and human antibiotic resistance rates (after controlling for contemporaneous changes in other factors, e.g., consumer awareness and education programs, HACCP interventions).
- *Genetic associations* between bacteria found in human patients and in food animals that may indicate whether they are similar enough so that one might come from the other (or whether both might have a common environmental source). Usually, epidemiological data are invoked to help interpret and complement genetic similarity data, since genetic similarities alone cannot establish a direction of causation.
- *Epidemiological associations* between exposures to food animal products and incidence rates of foodborne illnesses and/or prevalence rates of resistance in patients, after controlling for potential confounders, information biases, and modeling biases.

Well-developed statistical methods and algorithms are available to identify significant statistical associations from such relevant data (e.g., Mather *et al.*, 2004) and to screen them for potential causality based on the above criteria.

Technical Note: Statistical tests and algorithms for assessing potential causality. As mentioned earlier, statistical methods are available to identify exposure-response associations that are potentially causal, in that they cannot be "explained away" by conditioning on any other variables, even in very large data sets (Aliferis, 2003). The following information-theoretic criteria are useful for identifying evidence of potential causality in epidemiological data that may contain nonmonotonic or threshold-type dose-response nonlinearities. Roughly, a data set provides evidence that exposure variable X is a *potential cause* of response variable Y if and only if: (a) X is INFORMATIVE about Y, i.e., the mutual information between X and Y, denoted by I(X; Y) and measured in bits, is positive in the data set. (This provides the required generalization to not-necessarily-monotonic relations of statistical association measures for monotonic relations.) As a practical algorithmic matter, this implies that X should appear as a split in classification trees for Y, and *vice versa*; see Lemon *et al.*, 2003. [Technically, uncertainty about any discrete random variable X taking values x_i with corresponding probabilities p_i can be quantified by its *entropy*, defined as: H(X) = entropy of $X = -\Sigma_i p_i \log_2 p_i = E[\log_2(1/p_i)]$ bits. H(X) may be interpreted as the average amount of information gained when the value of X is learned. The *mutual information* between two random variables X and Y, denoted I(Y ; X), is defined as: $I(Y ; X) = H(Y) - H(Y \mid X)$ where $H(Y \mid X) = \Sigma_X Pr(X = x)H(Y \mid X = x) = E_X[H(Y \mid X)]$ is the conditional entropy of Y given X. For any specific observed value of X, say, x, the conditional entropy of Y given that X = x is: $H(Y \mid X = x) = -\Sigma_i Pr(Y = y_i \mid X = x)\log_2 Pr(Y = y_i \mid X = x)$. Classification tree algorithms grow trees recursively by starting with the dependent variable and always

"splitting" (i.e., conditioning) any currently unsplit node on the variable having the greatest estimated mutual information with the dependent variable, starting from the unsplit node, i.e., given the splits already performed for the set of cases described by it. The process ends when no statistically significant additional splits can be discovered. The tree many then be pruned back to minimize estimated true errors using cross-validation.] (b) UNCONFOUNDED: X provides information about Y that cannot be removed by conditioning on other variables, i.e., $I(X ; Y \mid Z) > 0$ for all subsets of variables Z. (Thus, splitting on Z first does not prevent X from entering the classification tree for Y.) (c) PREDICTIVE: Past values of X are informative about future values of Y, even after conditioning on past values of Y, i.e., $I(X^{-}(t) ; Y^{+}(t) \mid Y^{-}(t)) > 0$, where $X^{-}(t)$ denotes the set of X values at times $\leq t$, $Y^{-}(t)$ the set of Y values at times $\leq t$, and $Y^{+}(t)$ the set of Y values after t. [This generalizes the concept of Granger causality for multiple time series (e.g., Guatama and Van Hulle, 2003.)] (d) CAUSALLY ORDERED: Y is conditionally independent of the parents of X, given X, i.e., $I(P ; Y \mid X) = 0$, for any parent or ancestor P of X.. These criteria yield practical algorithms for detecting *evidence of potential causation* in cohort, case-control, and time series data sets. (Causation may be present even if these conditions are not satisfied, but then the data do not provide evidence of it.) These algorithms typically require tests for conditional independence as sub-routines. Classification tree software (Lemon *et al.*, 2003) can be used to perform conditional independence tests for one dependent variable at a time by testing whether conditional mutual information is significantly different from zero (Friedman, 1996; Frey *et al.*, 2003). Alternatively, statistical tests of the residuals in flexible nonparametric ("form-free") regression models (Shipley, 2000) can be used to test conditional independence for one dependent variable at a time. More computationally-intensive commercial software (e.g., BayesiaLab™) will automatically compute conditional independence relations for entire sets of variables (Tsamardinos *et al.*, 2003). These algorithms generalize the requirement that, to be considered causal, an exposure-response association must *not* be fully explained by confounding (Sonis, 1998; Greenland and Morgenstern, 2001; Greenland, 2003) – or, for that matter, by sample selection biases (Mark, 1997), information biases (Grimes and Shulz, 2002), or modeling and analysis biases (Cox, 2001). Formal tests for statistically significant associations between the timing of one event (e.g., introduction or cessation of animal antibiotic use) and subsequent changes in a series of measurements (e.g., human antibiotic resistance rates in a surveillance program) can be based on *intervention analysis* and *change point analyses* (Green, 1995) for time series. Potential causality between two time series of measurements (e.g., usage levels of an animal drug and illness rates or resistance rates in human patients) can be based on extensions of *Granger-Sims tests* (Swanson *et al.*, 2001) that include conditional independence and causal graph tests. These methods represent the current state-of-the-art in testing for potential causality. They are entering common biostatistical and risk analysis practice only slowly, but have been developed for many decades in other disciplines (Shipley, 2000).

Example: Tracking Sources of Resistant *E. faecium* by Phenotype

It is an *a priori* plausible hypothesis that antibiotic-resistant *E. faecium* isolated from human patients might originate in antibiotic-treated food animals that carry *E. faecium* (Wegener *et al.*, 1999). The hazard identification step of the risk assessment process rigorously tests and evaluates such hypotheses using data. It may also use epidemiological, time series, genotype, phenotypic biomarker, and other mechanistic data to investigate sources of exposures to microbial hazards (e.g., antibiotic-resistant *E. faecium*) even without any specific *a priori* hypotheses.

How can microbiological hazard identification methods be applied to identify sources of antibiotic-resistant *E. faecium*? One approach is to compare resistance phenotypes in isolates from different sources. This is illustrated by a study of Iversen *et al.* (2004):

> "An ampicillin- and ciprofloxacin-resistant *Enterococcus faecium* (ARE) strain, named FMSE1, with a characteristic biochemical phenotype, was in a recent study found to dominate among faecal ARE isolates from patients in several Swedish hospitals. In the present study, the prevalence of this strain among 9676 enterococcal isolates from healthy children, hospital sewage, urban sewage, surface water, slaughtered animals (broilers, pigs and cattle) and pig faeces and manure was investigated. Enterococcal isolates having the same biochemical phenotype as the FMSE1 were most common in samples of hospital sewage (50%), surface water (35%), treated sewage (28%) and untreated sewage (17%), but rare in samples from healthy children (0.8%) and animals (2%). PFGE typing of FMSE1-like isolates from hospital sewage indicated that they were closely related to the nosocomial FMSE1 strain. Thus, this study indicated a possible transmission route for nosocomial *E. faecium* from patients in hospitals to hospital sewage and urban sewage, and further via treatment plants to surface water and possibly back to humans. This proposed route of circulation of drug-resistant enterococci might be further amplified by antibiotic usage in human medicine. In contrast, such transmission from food animals seems to play a negligible role in Sweden."

Such studies can help to identify hazards that were not necessarily expected *a priori,* such as resistant bacteria in sewage and surface water. Conversely, they can help show the extent to which potential hazards that were expected *a priori*, such as the food animal transmission pathway, contribute to human illness in reality. The conclusion from this study, that transmission from food animals seems to play a negligible role in Sweden, might not have been anticipated in the absence of data-driven hazard identification, as many scientists have taken as axiomatic the assumption that foodborne transmission plays a major or predominant role in human resistant illnesses (e.g., APUA, 2002; Wegener, 2003).

The following example illustrates a very different use of biological information in hazard identification for antimicrobial resistance risk assessment. Comparing the rates and understanding the molecular biological mechanisms by which specific bacteria acquire resistance to specific antibiotics can indicate whether a hypothesized flow of resistance is plausible from the standpoint of the level of the organisms and mechanisms involved, without requiring population-level epidemiological data.

Example: Antibiotic Resistance Transfer Rates in *E. faecium* From Pigs

Use of virginiamycin in pigs may select for *E. faecium* that are both streptogramin-resistant and macrolide resistant (Aarestrup *et al.*, 2001). Jensen *et al.* (2002) identified vat(E) and erm(B) resistance genes in *E. faecium* isolates from pigs in Denmark:

> "The genetic background for streptogramin resistance was examined in *Enterococcus faecium* isolated from pigs (n = 55) and broilers (n = 207) in 1997 in Denmark. Fifty-one percent and 67%, respectively, of the isolates were resistant to streptogramins. Among streptogramin-resistant *E. faecium* (SREF), the genetic background for streptogramin A resistance could be determined in 96% of the isolates from broilers, compared with 14% among SREF from pigs. For broiler isolates 89% of SREF contained the *vat(E)* gene and 10% the *vat(D)* gene. Three of these isolates contained both resistance genes. **Among SREF from pigs two isolates contained the *vat(E)* gene and two others the *vat(D)* gene.** The genetic background for streptogramin B [resistance] was most often identified as the *erm(B)* gene encoding macrolide, lincosamide, and streptogramin B (MLSB) resistance. **Among SREF, 84% and 86% of isolates from broilers and pigs, respectively, contained the *erm(B)*.** In SREF from broilers, the *erm(B)* gene was physically linked to the *vat(E)* gene in 62% of the *vat(E)*-positive isolates and 79% of the isolates *containing vat(D). erm(A)* was detected in two SREF of broiler origin. Both isolates also contained the *erm(B)* gene." (Jensen *et al.*, 2002, emphases added)

Simjee *et al.* (2002) found that resistance to erythromycin, tetracycline, and streptomycin could be transferred *in vitro* from poultry-derived *E. faecalis* to *E. faceium* via conjugation and transmission of a plasmid containing the vat(E) gene.

However, analogous *in vivo* transfer from pig to human bacteria under real conditions is not known to occur. To the contrary, as noted by Aarestrup *et al.* (1997), "More resistance to streptomycin was observed among *C. coli* isolates from swine (48%) than among *C. coli* isolates from broilers (6%) or humans (0%)". Similarly, transfer of vancomycin resistance from food animals to human patients via *E. faecium* appears to be rare or non-existent. For example, in Denmark, Aarestrup *et al.* (2000) found that "All *E. faecium* isolates from humans were susceptible to vancomycin, whereas 10% and 17% of isolates from broilers and pigs, respectively, were resistant", suggesting that the flow of vancomycin resistance from

food animals to people is not very common. Since streptogramin resistance causes no clinical adverse consequences in human patients unless vancomycin resistance is also present (i.e., a streptogramin drug such as quinupristin-dalfopristin is prescribed only if vancomycin is not effective), the fact that vancomycin resistance is not readily transferred from swine to humans via the food chain also limits the potential human health risk from transfer of streptogramin-resistant *E. faecium*. Such data can thus help to inform human hazard identification for streptogramin use in pigs.

Example: A Caveat in Interpreting Resistance Rate Data

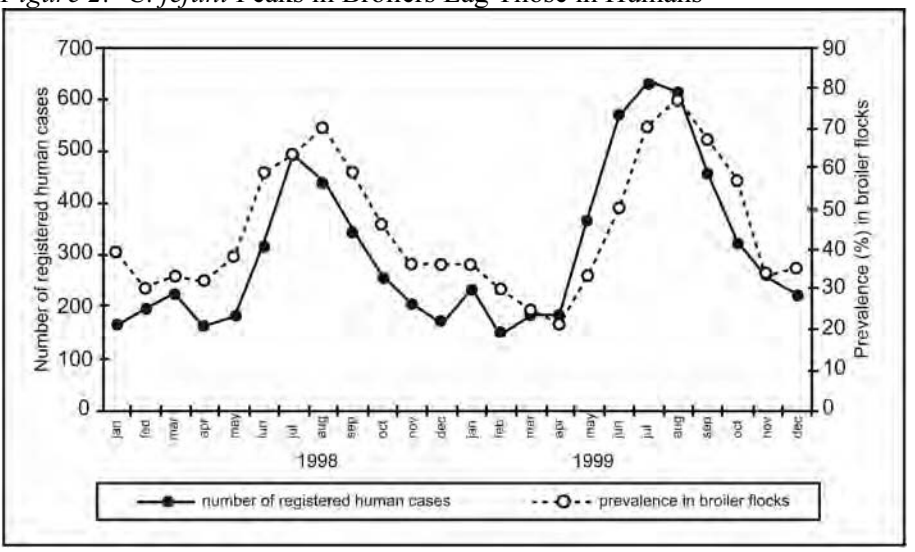
Problem: Suppose that a study of resistance rates to an animal antibiotic A in many different times and locations reveals that isolates of bacterium B from humans show a significant increase in the proportion of isolates resistant to A in just those locations where A has been used as an animal antibiotic, and in no other locations. Moreover, the increase in resistance rates always occurs immediately after antibiotic A is introduced for use in animals, and never before. Do these facts suggest that use of A causes an increase in resistance rates and resistance risks in humans? (Assume sample sizes are large enough so that these patterns are statistically significant.)

Solution: These facts do suggest a causal relation between introduction of animal antibiotic A and increases in the resistance rates, i.e., in the proportions of human isolates of B that are resistant to A. However, they do *not* necessarily suggest that use of A in animals increases the risk of resistant cases in humans (i.e., the expected number of resistant cases per year in the population, or per capita-year for any member of the population.) As a counterexample, suppose that the only effect of antibiotic A is to kill *all* susceptible strains of bacteria B and to kill most, but not all, of the A-resistant strains of B as well. Then the *proportion* of resistant isolates in humans infected with B bacteria from animals would increase, simply because the proportion of susceptible isolates has decreased. But *risks* (i.e., expected cases per year) of both susceptible and resistant infections would be decreased by the introduction of animal antibiotic use of A (Schwaber *et al.*, 2004).

Example: Using Seasonal Data in Hazard Identification

This example illustrates one way to sharpen and apply the traditional criterion of "temporality". In their quantitative risk assessment of *Campylobacter jejuni* in chicken products, Christensen *et al.* (2001) show the comparison of time series in Figure 2. They note that "It has been stated that the broiler flocks tend to peak after the human cases (Kapperud *et al.*, 1993; Berndtson, 1996; Newell *et al.*, 1999). This tendency is also seen in Denmark (Fig. 9). However, broilers seem to be infected before humans in 1998 and vice versa in 1999. If poultry are the primary source of human infection, it should be expected that the broilers peak before or coincident with the humans and not the other way around."

Figure 2: C. jejuni Peaks in Broilers Lag Those in Humans



The number of registered human *Campylobacter* cases and the *Campylobacter* prevalence in broiler flocks in 1998 and 1999 (Danish Veterinary Laboratory and Statens Serum Institut, unpublished data).

Source: Christensen *et al.*, 2001

This example illustrates one basic way to use time series information to test whether a causal hypothesis is consistent with data: by examining whether the hypothesized cause precedes the hypothesized effect. If not, as in the cited data sets, then the possibility of *reverse causation* (e.g., *Campylobacter* from human sources flowing to chicken flocks) or of a *common cause* (e.g., both chickens and human being exposed to *Campylobacter* carried via surface water, flies, etc.) should be considered. Genetic data can potentially help to clarify whether chickens, humans, and other warm-blooded animals tend to get the same strains of *Campylobacter* at the same time, or whether occurrence in some species (or other reservoirs) tends to precede occurrence in others.

Example: Retrospective Impact Assessment of Interventions

A second basic way in which time series information can be used in hazard identification is to test whether changes in exposures to a hypothesized hazard are followed by the changes in effects that would be expected if the hypothesized causal relation between exposure and effects were correct. For example, it has been widely hypothesized that use of animal antibiotics, including the fluoroquinolone drug enrofloxacin in chickens, is largely responsible for emergence of resistance to fluoroquinolones (such as ciprofloxacin) among human campylobacteriosis patients (e.g., FDA, 2001; Iovine and Blaser, 2004). On this basis, regulators have proposed

to withdraw approval for continued uses of enrofloxacin and other animal antibiotics, believing that doing so will help to preserve their efficacy in human medicine by reducing resistant cases of campylobacteriosis and other illnesses in humans (*ibid*). This belief need not remain an article of faith: it is empirically testable in countries where animal antibiotic uses have already been discontinued.

Comparing the time series of human antibiotic resistance rates for several drugs (including fluoroquinolones) before and after animal growth promoter bans in different countries shows some disappointing *increases* of resistance in humans and poultry following bans (e.g., Desmonts *et al.*, 2004 for France; Luber *et al.*, 2003 for Germany; Hayes and Jensen, 2003, for Denmark; Casewell *et al.*, 2003 for the European Union). Formal statistical analyses (e.g., interrupted time series models, intervention analyses, and evaluation studies) can help to quantify whether and to what extent these increases are different from what would have been expected in the absence of intervention (e.g., if bans had not occurred.) But even without such formal analysis, it is plain that causal hypothesis-based predictions that discontinuing animal antibiotic use would soon reduce human antibiotic resistance levels turned out to be incorrect.

Paying attention to such empirical evidence may help other countries to learn from European experience and to develop improved causal models, hazard identifications, and risk assessments as a basis for action, potentially leading to better public health outcomes. Formal hazard identification and risk assessment were not used as a basis for decision-making about the bans in the European Union. A "Precautionary Principle" approach was used instead, based on common-sense perceptions and expert judgments about cause-and-effect and recommendations for action (Phillips *et al.*, 2004). A problem with substituting precautionary principles and judgments for risk analysis informed by thorough understanding and testing of cause-and-effect relations is a version of the "law of unintended consequences": that interventions may not achieve their desired ends (and may bring about undesired ones) if they are not based on sound causal models.

The following example illustrates a situation in which historical data from different countries can be used to examine the human impacts of *not* intervening to halt the spread of antibiotic resistance. In this case, the evidence appears reassuring, i.e., macrolide resistance levels in humans have generally remained fairly low and stable, despite continued use of macrolides in food animals. Similar patterns hold for streptogramins. Such observations tend to falsify the testable prediction that continued use of animal antibiotics in food animals will rapidly increase human resistance levels. Whether historical stability of human resistance rates should be cause for complacency is less clear; this topic is addressed in Chapter 7.

Example: Retrospective Impact Assessment of Non-Intervention: Historical Macrolide Resistance Trends in Humans

Resistance to macrolides (e.g., erythromycin) in *Campylobacter* isolates from humans has remained relatively low in multiple countries over several decades, even though resistance to fluoroquinolones had increased rapidly (especially in Spain and parts of Asia). For macrolides, relevant historical evidence from various countries includes the following (all emphases added):

- *Canada* (Gaudreau and Gilbert, 2003): "The rates of resistance of 51 to 72 human strains of *Campylobacter jejuni* subsp. *jejuni* isolated annually from 1998 to 2001 in Montreal, Quebec, Canada, varied from 1 to 12% for erythromycin, 43 to 68% for tetracycline, and 10 to 47% for ciprofloxacin. In the last years of the study, **there was a significant increase in the rate of resistance to ciprofloxacin (P = 0.00003) but not in the rate of resistance to erythromycin** (P = 0.056) or tetracycline (P = 0.095) compared to the rate obtained in the first years."
- *Finland* (Rautelin *et al.*, 1991): "The results indicated that **susceptibility to erythromycin**, gentamicin, and doxycycline **has remained the same during the past 10 years**. … Resistance to erythromycin was 3% in both groups of strains. However, the number of norfloxacin-resistant strains increased from 4 to 11% in the follow-up period, and ciprofloxacin-resistant strains, which had not occurred 10 years ago, composed 9% of the strains isolated in 1990." The Finnish experience is updated in Rautelin *et al.*, 2003, which reports that "The growth of all domestic isolates was inhibited by erythromycin at concentrations of 4 microg/ml. … For the foreign isolates, the erythromycin MIC(90) was still low (4 microg/ml), but their susceptibilities to fluoroquinolones were clearly reduced."
- *Norway* (Afset and Maeland, 2001): "We examined retrospectively the *in vitro* susceptibility to erythromycin and ciprofloxacin of 296 *C. jejuni* strains isolated during the 1998-99 period. … Only one isolate showed resistance to erythromycin. …**Resistance of *C. jejuni* to erythromycin occurred very rarely**."
- *Spain* (Sanchez *et al.*, 1994): "We studied the evolution of antimicrobial susceptibilities of 275 clinical isolates of microorganisms of the genus *Campylobacter* isolated in our institution during a 5-year period (1988 to 1992). The microorganisms studied were *C. jejuni* (n = 230), *C. coli* (n = 42), and C. *fetus* (n = 3). The overall resistance rates (determined by the agar dilution method and the recommendations of the National Committee for Clinical Laboratory Standards) were as follows: erythromycin, 2.3%; clarithromycin, 2.3%; azithromycin, 1.9%; ciprofloxacin, 28.5%; norfloxacin, 31%; ofloxacin, 26.3%; and nalidixic acid, 36.8%. The evolution of resistance (percent resistance in 1988 versus percent resistance in 1992) was as follows: erythromycin, 2.6 versus 3.1; clarithromycin, 2.6 versus 3.1; azithromycin, 2.6 versus 3.1; ciprofloxacin, 0 versus 49.5; norfloxacin, 2.6 versus 55.5; ofloxacin, 0 versus 45.6; nalidixic acid, 2.6 versus 56.8. **Our data show stable macrolide activity against *Campylobacter* spp.** and the rapid development of quinolone resistance over the last 5 years."
- *Sweden* (Osterlund *et al.*, 2003): "Antibiotic resistance was compared in 844 *Campylobacter* jejuni/coli strains acquired outside Sweden and 575 acquired in

Sweden during 1990-2002. There was a **clear gradual increase in ciprofloxacin and tetracycline resistance among C. jejuni/coli strains acquired outside Sweden during the 13 y period. This trend was not seen for erythromycin** or in domestically acquired strains for any of the 3 antibiotics tested." In earlier decades, it was found that, from 1978-88, "**No general increase in *in vitro* resistance to antibiotics commonly used for the treatment of human gastroenteritis caused by *C. jejuni* or *C. coli* has occurred during the last 10 years in Sweden**, which might be a consequence of strict antibiotic control. The numbers of strains from 1988 to 1989 resistant to ciprofloxacin and to norfloxacin included in this study (0.7 and 1.4%, respectively) are still fewer than those that were resistant to erythromycin (7.3%) or doxycycline (12.4%). There is, however, since 1989 to 1990 an indication of increasing resistance to these [fluoroquinolone] antibiotics." (Sjogren *et al.*, 1992). (From 1992 to 1995, it appeared that "The MIC50 and MIC90 values for doxycycline and erythromycin have increased markedly through the 4 years studied." (Sjogren *et al.*, 1997), although Osterlund *et al.*, 2003 found no long-term increase in the trend for erythromycin from 1990-2002)

- *United States:* CDC-NARMS monitoring data using the E-test show that **erythromycin resistance in *C. jejuni* declined from 8% in 1997 to 3% in 1998, 2% in 1999, and 1.3% in 2000** (www.cdc.gov/narms/annual/2000/annual_pdf.htm). (These data are suspect, i.e., a pattern of steady decline over these years may not be correct, but neither was there an increase.)

Germany may be an exception, however. Luber *et al.*, 2003, report that: "Among human *C. coli* strains the rate of resistance to erythromycin rose from 7.1% in 1991 to 29.4% in 2001-2002. In comparison, *Campylobacter* sp. isolates from poultry already had high rates of resistance in 1991. …Thus, discrepancies in the antimicrobial resistance rates among *Campylobacter* isolates originating from poultry and humans support the hypothesis that at least some of the resistant *Campylobacter* strains causing infection in humans come from sources other than poultry products."

In summary, although resistance to macrolides in human patients may increase over time in some geographic areas, especially if human use is high, the overall pattern of international trends has exhibited relatively low levels of macrolide resistance in humans in most countries over the past decades, and there appears to be little evidence of a relation between continued use in animals and increasing resistance levels in human patients.

Example: Caveats for *Ex Post* Interrupted Time Series Analysis

Vellinga and Van Look (2002) reported a fascinating use of an unplanned interruption in chicken consumption in Belgium to test whether campylobacteriosis rates declined when chicken consumption decreased. They report: "In June 1999, the dioxin crisis, caused by dioxin-contaminated feed components, exploded in Belgium, resulting in withdrawal of chicken and eggs from the market. Through the

sentinel surveillance system, a decrease in *Campylobacter* infections during June 1999 was noticed. A model was generated with the reports from preceding years (1994 to 1998), and a prediction of the number of infections in 1999 was calculated. The model shows a significant decline (40%) in the number of infections, mainly because of the withdrawal of poultry. The use of a disaster as an epidemiologic tool offers a unique opportunity to observe exceptional changes in the occurrence of infections or other diseases." This example illustrates a potentially valuable and valid approach to providing positive evidence that a causal hypothesis is correct.

Unfortunately, in this particular study, the authors did not use standard methods of interrupted time series analysis to analyze the data, and their conclusions do not appear to be valid. The time series of weekly *Campylobacter* counts shows enough variability in each year so that the claimed 40% decrease in the weeks following the withdrawal is not remarkable or unexpected. The analysis does not demonstrate that the time series behaved significantly differently in June of 1999 compared to June of other years. [In addition, there may be data integrity issues, e.g., Vellinga and Van Look report a dramatic *decrease* from counts in May (weeks 18-21) to counts in June (weeks 22-25), which appears to be inconsistent with data showing an *increase* over this interval published by the Belgian Institute of Public Health (http://www.iph.fgov.be/epidemio/epien/index0000.htm). The duration of the ban cited by Vellinga and Van Look also appears to have been mistakenly increased from about 2 to about 4 weeks.] Despite these difficulties in the detailed analysis and conclusions, the basic idea is sound and important: a causal hypothesis, such as that chicken consumption causes a large share of campylobacteriosis, can be tested empirically in situations where exposure is interrupted or otherwise strongly perturbed. However, in such cases, it is best to present the results of standard methods of analysis for interruptions, interventions, or change points in time series, so as to avoid errors and confusions in the interpretation of the results.

In summary, although it may be difficult or impossible to *prove* causation from data, it usually *is* possible to test, and perhaps refute (if they are mistaken) competing non-causal hypotheses; and also to test whether the hypothesized causal relations between decisions and exposures and between exposures and risk are *consistent* with available data (Shipley, 2000). Time series data, cross-sectional (e.g., case-control) data, and biological data on microbial phenotypes, genotypes, and resistance rates and mechanisms, afford a number of opportunities to confirm or refute testable causal hypotheses and predictions. Candid presentation and discussion of the results of such tests and of their implications for causal inferences and for the value of further information that can resolve ambiguities in causal theories are valuable outputs of a successful hazard identification.

The preceding examples have provided relatively simple illustrations of principles and methods useful in hazard identification. In practice,

however, it may be necessary to combine several techniques. Even then, the results may be ambiguous. The following case study illustrates some of the practical complexities of hazard identification with partly conflicting, less than decisive, data.

5. CASE STUDY: DID ENDING AVOPARCIN USE REDUCE HUMAN ANTIBIOTIC RESISTANCE?

Figure 1 of Chapter 1 showed that, following a ban on macrolides (tylosin) and other growth promoters in Denmark in 1998, macrolide (erythromycin) resistance in *C. jejuni* among domestically acquired human cases promptly rose in 1999, again in 2000, and still further in 2001 (Hayes and Jensen 2003). WHO (2003, p. 28) noted that, in Denmark, "There is also an indication that since the termination of growth promoters there may have been an increase in resistance among *E. faecalis* to the macrolide drug erythromycin" and remarked (p. 6) that "Direct effects of the termination of growth promoters on resistance in gram-negative bacteria (e.g., *E. coli*, *Salmonella*) were neither expected nor observed." For fluoroquinolones (enrofloxacin), too, as mentioned previously, reductions in use in different countries have been followed by increases in fluoroquinolone resistance among both human and animal isolates (e.g., Desmonts *et al.*, 2004; Luber *et al.*, 2003). Such empirical observations raise an obvious question for hazard identification: *Does withdrawing animal antibiotics actually cause reductions in antibiotic resistance to similar antibiotics among human patients?* The answer is central to the animal antibiotic debate.

In general, whether and to what extent terminating antibiotic uses in food animals would cause a quick decline (or any decline) in resistance in bacteria in either animals or people in any given country is an open question. The epidemiology of resistance to major classes of antibiotics in human and animal populations differs significantly across countries and continents, as revealed by studies of vancomycin-resistant enterococci (VRE) in Europe, the US, Australia, and elsewhere (e.g. Goossens, 1998; Bell *et al.*, 1998).

European experience following withdrawals of the animal antibiotic avoparcin (which selects for VRE, since vancomycin and avoparcin are both glycopeptides) has been mixed. At first, some investigators claimed that in Germany, as hoped, VRE rates were declining rapidly and significantly in food animals and among healthy humans in the community (Klare *et al.*, 1999; Witte, 2000). But further data soon raised anomalies. In Germany, other investigators had already noticed that the resistance patterns of the VRE isolates obtained from food (including minced pork) were different from the resistance patterns of clinical isolates from humans, suggesting that the food chain might not, after all, be responsible for the incidence of highly

resistant (vanA) VRE observed in human nosocomial infections (Klein *et al.*, 1998). This was soon confirmed by additional observation (Lemck and Bulte, 2000).

In Italy, Del Grosso *et al.*, 2000 reported that "After the avoparcin ban, a decrease in the rate of VRE contamination in meat products was observed. Such a decrease was statistically significant in poultry (from 18.8% to 9.6%) but not in pork products (from 9.7% to 6.9%). The majority of VRE from all sources carried the vanA resistance gene and included *Enterococcus faecium, E. faecalis, E. hirae, E. durans,* and *E. gallinarum.* None of the strains carried the vanB gene, whereas constitutively resistant vanC-positive strains were frequently found. Our results show that avoparcin withdrawal has been successful in reducing VRE contamination in meat products."

In Norway, the highly resistant vanA VRE showed unexpected persistence on poultry farms. Even five years after the avoparcin ban, it was found that "VRE were isolated from 72 of 73 (99%) and eight of 74 (11%) poultry samples from exposed and unexposed farms, respectively. VRE were isolated from 13 of 73 (18%) and one of 74 (1%) farmer samples from exposed and unexposed farms, respectively. All VRE isolates were highly resistant to vancomycin and possessed the vanA gene, as shown by PCR. The high prevalence of VRE is in accordance with previous Norwegian studies, and shows a remarkable stability of the VanA resistance determinant in an apparently non-selective environment" (Borgen *et al.*, 2000).

In Denmark, a prolific and influential group of researchers had long argued that animal antibiotic use could significantly increase the rates of resistance in human bacteria and had advocated banning animal growth promoters to reduce resistance levels in human patients. They later published data and interpretations suggesting that the bans were highly effective in reducing resistance in food animals and people. For example:

- Bager *et al.* (1999) stated that "Among isolates from broilers, the proportion that were resistant to glycopeptides [such as vancomycin] has shown a statistically highly significant decline between the end of 1995 and the first half of 1998, whereas in pigs the ban appears to have no such effect."
- Aarestrup *et al.* (2001) announced that "The avoparcin ban in 1995 was followed by a decrease in the occurrence of glycopeptide-resistant *E. faecium* (GRE) in broilers, from 72.7% in 1995 to 5.8% in 2000."
- Wegener, 2003, citing Aarestrup *et al.* (2001) and other sources, claimed that "The data shows that although the levels of resistance in animals and food, and consequently in humans, has been markedly reduced after the termination of use [of avoparcin], the effects on animal health and productivity have been very minor."

However, other investigators have not independently reproduced these conclusions. For example, Heuer *et al.*, 2002 found "no significant decrease in the proportion of VRE-positive [broiler] flocks during the study period (1998-2001)" and concluded that "This study demonstrated the extensive occurrence of VRE in broiler flocks more than 5 years after the avoparcin ban in Denmark, and indicates that VRE may persist in the absence of the selective pressure exerted by avoparcin. The results differ markedly from previously published Danish surveillance data on VRE in broilers. This may reflect differences in isolation procedures." More recently, persistence of VRE after withdrawal of avoparcin has also been documented in New Zealand (Manson *et al.*, 2004), indicating that it is not a Europe-specific phenomenon.

In Sweden, avoparcin use halted in 1986. However, Iversen *et al.* (2002) found that "Surprisingly, VRE were isolated from 21 of 35 untreated sewage samples (60%), from 5 of 14 hospital sewage samples (36%), from 6 of 32 treated sewage samples (19%), and from 1 of 37 surface water samples. … Most of the VRE were multiresistant. …We conclude that VRE [mostly vanA vancomycin-resistant *E. faecium*] were commonly found in sewage samples in Sweden. The origin might be both healthy individuals and individuals in hospitals. Possibly, antimicrobial drugs or chemicals released into the sewage system may sustain VRE in the system." Thus, it appears that high levels of VRE, specifically including vanA resistance in *E. faecium*, are still frequently found in sewage more than fifteen years after avoparcin stopped being used in animals.

Finally, whether or not animal antibiotic use contributes to the selection and maintenance of VRE in pigs, the relevance of such VREs to human antibiotic resistance and health impacts is unclear. For example, Willems *et al.*, 2001 concluded that hospital epidemics of vancomycin-resistant *E. faecium* (VREF) on three continents are caused by a genetically distinct subpopulation (carrying the *esp* gene) rarely found in animals or healthy members of the community:

> "In the USA, vancomycin-resistant *Enterococcus faecium* (VREF) is endemic in hospitals, despite lack of carriage among healthy individuals. In Europe, however, hospital outbreaks are rare, but VREF carriage among healthy individuals and livestock is common. We used amplified fragment-length polymorphism analysis to genotype 120 VREF isolates associated with hospital outbreaks and 45 non-epidemic isolates from the USA, Europe, and Australia. We also looked for the *esp* virulence gene in these isolates and in 98 VREF from animals. A specific *E. faecium* subpopulation genetically distinct from non-epidemic VREF isolates was found to be the cause of the hospital

epidemics in all three continents. This subpopulation contained a variant of the *esp* gene that was absent in all non-epidemic and animal isolates."

In summary, the temporal relation between withdrawal of avoparcin and subsequent VRE and VREF levels in animals, healthy people, and hospitalized patients is at best ambiguous and at worst may be absent. While some investigators have announced clear temporal relations between cessation of animal antibiotic use and subsequent reductions in resistance rates in food animals and/or people, others have not found any such decline and instead have commented on the unexpected persistent of resistance years, and even decades, after withdrawal of the animal antibiotics.

The mixed evidence in this case about effects of avoparcin withdrawal on glycoprotein antibiotic resistance among bacteria from animals, meats, member of the community, and hospitalized patients highlights the importance of distinguishing among these endpoints (only the last of which ultimately matters for hazard identification for VREF) and between the fraction of human health impacts per year that are *attributed* to past animal drug use and the fraction that would disappear, i.e., would be *prevented*, if that drug use were to cease. It is important not to conflate them. The following two questions about causality are distinct: (1) did historical use of an animal antibiotic (such as enrofloxacin, tylosin, or avoparcin) *cause* (i.e., hasten, contribute to or bring about) increased levels of resistant bacteria in human patients (e.g., VRE infections or illnesses from resistant strains of *C. jejuni*, *C. coli*, or *Salmonella*)? *vs.* (2) would halting the animal antibiotic use now *prevent* or reduce future levels of resistant bacteria? To inform current risk management decisions, the latter is the more relevant question, and the one that hazard identification should help to address.

6. A SYSTEMS DYNAMICS APPROACH TO HAZARDS

On its face, the repeated finding that reducing animal antibiotic use is soon followed by increased resistance to the antibiotic among isolates from human patients and/or animals appears paradoxical. How can removing the selection pressure from an animal antibiotic fail to reduce the resistance levels in humans and in animals (or at worst, as for avoparcin and VRE, perhaps, leave them unchanged)? An *a priori* plausible conjecture might be that human antibiotic resistance rates increase with or without animal antibiotic use, e.g., because of human use of similar drugs. Terminating animal antibiotic use would then slow the increase (or have no effect, if it is entirely driven by other factors). But while this explanation might be plausible for some drugs, it appears to be inconsistent with data for others,

such as Figure 1 of Chapter 1, where resistance to erythromycin and to streptomycin among human *Campylobacter* isolates appears to increase specifically after the ban on macrolides and other growth promoters.

A systems dynamics approach may help to unravel the mystery. In dynamical systems theory, "causation" can be described straightforwardly in terms of changes in a system's outputs produced (i.e., caused) by changing its inputs and then applying the dynamic equations linking its inputs to its outputs. Hazard identification for antimicrobial risk assessments dealing with resistance risks can benefit by analyzing the qualitative behaviours of dynamical systems (or numerical simulations of their behaviours) to predict and describe the dynamic responses (e.g., outbreaks of antibiotic resistance epidemics, maintenance of stable endemic resistance levels, transient increases in resistance levels, etc.) caused by changing controllable inputs such as animal antibiotic use. This section illustrates the approach with a simple conceptual and mathematical model of possible dynamic interactions among animal and human illnesses and susceptible and resistant bacteria.

6.1 Model Definitions and Equations

A model of the human health impacts of animal antibiotic use should account for at least the following quantities:

- IH(t) = the fraction of the human population of interest that has a specified foodborne illness, such as campylobacteriosis or salmonellosis, at any time t. (IH = "ill humans" fraction.)
- IA(t) = the fraction of *servings* of a particular food commodity that comes from animals with a specified illness or adverse condition (e.g., airsacculitis or necrotic enteritis) that the animal antibiotic could help to prevent, reduce, or control. (IA = "ill animal" fraction for servings from processed animals. Animals not sent to slaughter and animal carcasses removed during processing are excluded from consideration when IA is calculated, as they presumably do not affect IH.)
- RH(t) = fraction of human infections that are resistant to the human antibiotic of interest at time t. (RH = "resistant human" fraction.)
- RA(t) = fraction of isolates from the food animals of concern (those that contribute to ingested servings of a meat commodity) that are resistant to the antibiotic of interest at time t.

For example, in a study of the human health impacts of enrofloxacin use in chickens, IH(t) might denote the fraction of the population with campylobacteriosis at time t, while RH(t) = fraction of those illnesses that are resistant to fluoroquinolones (ciprofloxacin), IA(t) = fraction of chicken servings from airsacculitis-positive flocks (a condition that can be treated

effectively with enrofloxacin), and RA(t) = fraction of isolates from chicken carcasses that are resistant to fluoroquinolones. In a study of the human health impacts of virginiamycin (VM) use in poultry, IH(t) might represent the fraction of immunocompromised intensive care unit (ICU) patients with $VREF_A$ (vanA vancomycin-resistant *E. faecium*) infections at time t; RH(t) might represent the fraction of these cases that are resistant to the human drug Synercid™ (quinupristin-dalfopristin, QD), the human analogue of VM; and IA(t) and RA(t) might represent the fraction of chicken servings from necrotic-enteritis positive flocks (controlled with VM) and the fraction of $VREF_A$ isolates from those servings that are QD-resistant, respectively. It is not necessary to define these quantities with great precision for purposes of hazard identification, as the main purpose of hazard identification is simply to identify qualitatively what human health impacts (i.e., what changes in IH(t) and RH(t)) are likely to be caused by changes in animal drug use. Thus, vexed questions such as the proper definition of "resistant" (from microbiological and clinical perspectives) need not be resolved before applying the model.

The four variables IA, RA, IH, and RH interact with each other over time. The dynamics of the model are governed by a system of ordinary differential equations (ODEs) describing how the levels of the quantities affect each others' rates of change. The controllable input to the system is the level of animal drug use. For simplicity, we will only consider the impacts of banning an existing animal drug use or of introducing a new one. We assume that the effect of the animal drug use is to promptly (within a few months, as one flock replaces another) reduce the fraction of ill animals going to slaughter from a higher level, IA_0, if no animal drug use is allowed, to a lower level, $IA_1 < IA_0$, if animal drug use is allowed. In addition, animal drug use exerts a selection pressure on animal bacteria that tends to increase RA. Define the decision variable as: A(t) = 1 if the animal drug is used at time t, else A(t) = 0. (A = "animal antibiotic use" or "action".) A(t) is the controllable input to the system, and the purpose of the model is to identify the potential human health consequences of changing it. Then the equations of a basic model are as follows:

$$IA(t) = A \times IA_1 + (1 - A) \times IA_0 \quad (1)$$

or, equivalently, $IA = IA_0 + A^*(IA_1 - IA_0)$. where short-run transients and within-flock dynamics of animal illnesses are ignored;

$$dIH/dt = [a_1 + b_1(1 - IA) + c_1IA + d_1IH](1 - IH) - r_1IH \quad (2)$$

$$dRA/dt = [a_2 + b_2A + c_2IA^*A + d_2IH + e_2RH](1 - RA) - r_2RA \quad (3)$$

$$dRH/dt = [a_3 + b_3A + c_3IA^*A + d_3IH + e_3RH + fRA](1–RH) - r_3RH \quad (4)$$

In these equations, the time dependency of the variables is not shown explicitly, i.e., we write A, IH, RA, and RH instead of A(t), IH(t), RA(t), and RH(t), respectively. Possible differences in virulence between resistant and susceptible strains of bacteria are not modeled, although the approach could be extended to include such generalizations.

The three ODEs have similar structures. In each, the rate at which some undesirable quantity X changes has the form:

$$dX(t)/dt = k[1 - X(t)] - rX(t), \text{ abbreviated as: } dX/dt = k(1 - X) - rX.$$

This template applies for X equal to each of IH, RA, and RH. The interpretation is that (1 – X) is proportional to the pool of "susceptibles": it is the fraction of units (people, animals, or bacteria) that do not currently have the undesirable condition (illness or resistance) but that may acquire it. k is the fractional rate per unit time at which susceptible units acquire the undesirable condition, while r is the "recovery rate" at which units with the undesirable condition make a transition back to not having it.

Specific expressions for the multiplier k in the generic template are given by the summed terms in square brackets on the right sides of equations (2)-(4). The coefficients, a_1, a_2, and a_3, represent the spontaneous transition rates from the susceptible to the affected groups (per susceptible unit, per unit of time). b_1 reflects the rate at which servings from healthy animals, represented by (1 – IA) (the fraction of animals that are not ill) make healthy people (represented by (1 – IH)) become sick; while c_1 is the analogous coefficient for servings from ill animals. If the disease is infectious, then $d_1 > 0$ represents the rate at which ill humans, IH, infect well ones, (1 – IH). If b_2 is positive, then the use of the animal antibiotic, e.g., as a prophylactic or growth promoter, contributes to the flow of animal bacteria from the pool susceptible to the antibiotic, (1 – RA), to the resistant pool, RA. b_3 allows for the possibility that such use may also contribute directly to selection of resistant strains in human, e.g., because use on the farm leads to runoff into surface water, selection of resistant strains in the environment, and hence an increase in the resistant fraction RH among isolates from affected humans. c_2 and c_3 allow for therapeutic uses of the animal antibiotic in ill animals to contribute to resistance selection in animals and human, respectively. d_2 and d_3 allow for treatment of ill humans to select for resistant strains that are then found in animals (as in Iversen *et al.*, 2004) and in humans, respectively, while e_2 and e_3 allow for the possibility that patients with resistant strains may excrete them and contribute differently to such strains in animals and humans, respectively, than do ill humans in general. f in equation (4) reflects the possibility that resistant bacteria are transferred from animals to humans, perhaps directly via food, or indirectly via transfer of resistance

determinants. For simplicity, we assume that the fraction of people with resistant illnesses at time t is simply IH(t)*RH(t), i.e., ill people are assumed to have the probability RH(t) of having resistant illnesses at time t. While other model structures can be envisioned, equations (1)-(4) suffice to illustrate the dynamic systems approach to hazard identification.

6.2 Model-Based Analysis of Potential Resistance Hazards

Collecting data from which to estimate all of these effect coefficients (or even the reduced set obtained by specifying that some of them are zero, that others are equal to each other, and then using algebra to combine like terms) could be a considerable undertaking and raise statistical challenges, such as whether the coefficients are uniquely identifiable from the available data. Fortunately, hazard identification does not require such a thorough approach. Rather, we care only about how changing the input A (from 1 to 0 for existing animal drug products, or from 0 to 1 for new products) would affect the levels of IH and RH. Methods for qualitative analysis of system behaviours suffice for this purpose.

Rather than considering all of the *a priori* possibilities expressed in the above model, it is useful to restrict attention to the following special case of it (in which d_1, b_3, c_2, e_2 and e_3 are all made into "structural zeros", reflecting an absence of these potential links):

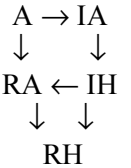
$$dIH/dt = [a_1 + b_1(1 - IA) + c_1IA](1 - IH) - r_1IH \quad (2')$$

$$dRA/dt = [a_2 + b_2A + d_2IH](1 - RA) - r_2RA \quad (3')$$

$$dRH/dt = [a_3 + d_3IH + fRA](1-RH) - r_3RH \quad (4')$$

This sub-model may be appropriate for an animal antibiotic used as a growth promoter or prophylactic (so that the usage rate is not confined to ill animals) and a bacterium for which the spread of resistant strains is driven by animal use (via b_2A) and by empiric treatment of ill humans (whether or not they have resistant strains, via d_2IH), perhaps followed by runoff of the prescribed human antibiotic to sewage and consequent selection and spread of resistant strains to animals and humans (via fRA) (Iversen *et al.*, 2004).

The dependencies among variables in this model are summarized in the following directed acyclic graph ("DAG" or "acyclic digraph"):



In this graph, an arrow from one variable pointing into another means that the first helps to determine the value of the second (i.e., the first appears on the right-hand side of the equation for the variable that it points into.) Each dynamic variable (with a time course described by an ODE) has an implicit "self-loop", i.e., its own current value, X, may help to determine its current rate of change, dX/dt, and hence its future values. The graph is acyclic only because we simplified the system of model equations (making it "recursive", in the terminology of econometrics). This allows an elementary analysis. Dynamics of systems with cyclic dependency graphs and with complicated dynamics, perhaps lacking stable steady-state equilibria, can be analyzed using more powerful mathematical methods of dynamical systems theory.

Equilibrium levels of human illness and resistance before and after a one-time change in the level of A can be studied by replacing the dynamic equation template $dX/dt = k(1 - X) - rX$ by the corresponding algebraic equilibrium condition (on setting $dX/dt = 0$):

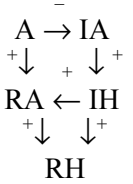
$k(1 - X) = rX$, which can be rearranged as: $X = k/(k + r)$

Clearly, any change that increases k will increase the steady-state value of X (by moving the ratio $k/(k + r)$ closer to 1, assuming $k > 0$ and $r > 0$.) The formulas for k for the three ODEs are:

- For IH: $k_1 = a_1 + b_1(1 - IA) + c_1IA = (a_1 + b_1) + (c_1 - b_1)IA$
- For RA: $k_2 = a_2 + b_2A + d_2IH$
- For RH: $k_3 = a_3 + d_3IH + fRA$

Thus, if $d_3 > 0$ and $f > 0$, then RH increases with both IH and RA. If $b_2 > 0$ and $d_2 > 0$, then RA increases with both IH and A. IA decreases with A, assuming that the animal antibiotic use is successful in reducing or preventing animal illnesses. Finally, IH increases with IA if and only if $c_1 > b_1$, for example, if servings from ill animals carry higher microbial loads and hence greater risk-per-serving than servings from well animals.

Assuming that these inequalities hold, the qualitative directions of the impacts on human health of a change in animal drug use can be assessed by assigning signs to the arrows in the dependency digraph, as follows.



Signed Digraph Model of Qualitative Impacts

By hypothesis, changing the value of input A from A = 1 to A = 0 (i.e., implementing a ban on an existing use) *increases* IA (from IA_1 to IA_0). However, it tends to decrease RA. Therefore, while the impact of the change in A on the equilibrium value of IH is an unambiguous increase (since reduced A implies increased IA and hence increased IH), the impact on RH depends on the relative strengths of the impacts from decreases in RA due to lower A (transmitted via the coefficients b_2 and f) and the impacts from increases in IH (transmitted via d_2, f, and d_3).

Elaboration: To simplify further, suppose that the effect of IH on RA is negligible ($d_2 \approx 0$). Then the explicit system of steady-state equilibrium equations of the form $X = k/(k + r)$ becomes:

- $IH = \{a + c[IA_0 + A^*(IA_1 - IA_0)]\}/\{r_1 + a + c[A^*IA_1 + (1 - A)IA_0]\}$
- $RA = (a_2 + b_2A)/(a_2 + b_2A + r_2)$
- $RH = (a_3 + d_3IH + f^*RA)/(r_3 + a_3 + d_3IH + fRA)$

where for brevity we have introduced the reduced coefficients $a = (a_1 + b_1)$ and $c = (c_1 - b_1)$ and we are now working with steady-state equilibrium levels. The impact of a change in A on RH is positive if and only if it increases the variable component of the numerator, $d_3IH + f^*RA$ (since this will increase the ratio defining RH, moving it closer to 1.) Changing A from 1 to 0 *reduces* RA by an amount ΔRA, from $(a_2 + b_2)/(a_2 + b_2 + r_2)$ to $a_2/(a_2 + r_2)$ and *increases* IH by an amount ΔIH from $(a + cIA_1)/(r_1 + a + cIA_1)$ to $(a + cIA_0)/(r_1 + a + cIA_0)$. Thus, the impact on RH is an increase if and only if $d_3\Delta IH > f^*\Delta RA$, i.e., $\Delta IH > (f/d_3)\Delta RA$.

The conclusion is that a ban in animal drug use can *increase* human antibiotic resistance levels, even while reducing animal resistance levels, if it leads to a sufficiently large increase in sick animals, and hence in sick people. This provides a possible theoretical explanation of the apparent paradox that human resistance levels to a drug can increase after animal use of the drug is banned. To test whether this explanation is consistent with data, it would be necessary to check whether the frequency of microbial loads of bacteria sufficient to cause illness does in fact increase after bans.

7. HAZARD IDENTIFICATION METHODS TO AVOID

This section provides some examples and brief discussions of common mistakes to avoid in hazard identification. As suggested by the preceding examples, valid hazard identification from data always requires care, and sometimes requires application of sophisticated statistical and causal modeling methods to avoid mistaken conclusions. Fortunately, identification of potential causes from multivariate data and rigorous testing of causal hypotheses for consistency with data are both increasingly well-supported by

current statistical methods and algorithms for causal analysis. However, these methods have historically been slow to enter the mainstream of biological and epidemiological thought and practice (e.g., Shipley, 2000).

Many current epidemiological papers and published conclusions about causation and hazards in microbial and antimicrobial risk assessment depend on less trustworthy methods, such as Hill-type criteria (which have no necessary relation to causation and can be misleading in the absence of careful refutation of competing explanations); traditional epidemiological "attributable risk" and "attributable fraction" calculations (which are based on statistical associations rather than on causation, and which also have no necessary relation to causal impacts); and common-sense judgments or precautionary assumptions, neither of which necessarily provides a reliable or useful guide to the truth about causation.

Common errors to avoid in hazard identification are as follows.

- *Non-causal hazard identification.* As discussed earlier in this chapter, simply identifying a microorganism in food without showing that it causes human health harm in practice does not constitute hazard identification. Even if it is known that the identified agent can be pathogenic under laboratory conditions, it is still necessary for a full hazard identification to demonstrate that it causes harm in reality. Likewise, hazard identification should usually not be based on interpretations of statistical regression coefficients as indicators of causal influences, since regression models typically deal with association rather than causation (Freedman, 2004).
- *Unspecified harm.* Identifying a pathogen without identifying any clinically relevant harm that it causes does not constitute adequate hazard identification. A hazard identification should demonstrate that exposure to the identified pathogen actually causes some specific adverse health effect, rather than merely that the agent is present. Although purely hypothetical or theoretical suppositions about harm caused by microorganisms or resistance determinants in food can be useful for hypothesis-generating purposes, hazard identification requires that the causal hypotheses be tested empirically, and that competing explanations be identified and refuted.
- *Partial or incomplete hazard identification.* Identifying only one component of hazard (e.g., effects of animal antimicrobial agents on human health effects of resistant bacteria but not susceptible bacteria) can give an incomplete description of potential risk that is not suitable for guiding rational decision-making.
- *Hazard identification based on temporal trends.* Discussions of hazard identification in antimicrobial risk assessment sometimes refer to temporal "trends" in which adverse health effects occurred *after* the

historical introduction of an antimicrobial agent in feeds, and are therefore suspected of being *caused* by it (e.g., FDA-CVM, 2001 for fluoroquinolone resistance in *Campylobacter*). Such discussions are usually inconclusive, neither establishing nor refuting the suggested causal explanation. In rigorous analyses of causation, such "temporal trend" arguments are generally dismissed as instances of *post hoc ergo propter hoc* (www.fallacyfiles.org/noncause.html) or "false cause" fallacies. If sufficient longitudinal data are available, they may be used in statistical tests for potential causation for hazard identification, as illustrated in the preceding section. However, these methods are based not on trends, but on verifying conditional independence relations and showing that significant changes in the data-generating process for a time series occur following interventions, such as introduction of a feed additive. Statistical methods of intervention analysis and change point analysis in time series can be used for this purpose.

- *Unjustified causal interpretation of statistical associations*. Occasionally, researchers report statistical associations as if they were causal relations. For example, Smith *et al.* (1999) noted that humans and chickens carry similar strains of fluoroquinolone-resistant campylobacter and concluded that "the number of quinolone-resistant infections acquired domestically has also increased, largely because of the acquisition of resistant strains from poultry". Yet, logically, identifying *similarities* in strains implies nothing about the *causes* of the similar strains (e.g., whether poultry are a source for humans, as asserted here; or human sewage is a source for poultry (reverse causation), or both come from a common cause, such as surface water.) (In this example, finding similar strains among other warm-blooded animals, such as lambs, would tend to strengthen the hypothesis of a common cause.)
- *Meaningless statistics and logical fallacies.* A fallacy that is surprisingly common in antimicrobial risk assessment is to compare data from two points in time that seem to suggest an increasing trend, even when the full data show no such trend. For example, suppose that an animal drug was introduced in 1997 and that an identical or similar human drug had been in use since 1990. If human resistance rates for the drug in different years are as follows: 12% in 1994; 10% in 1995; 8% in 1996; 12% in 1997; 10% in 1998; 8% in 1999; 12% in 2000; and 10% in 2001, then it would clearly be a fallacy (a form of "cherry picking") to compare the resistance rates in 1996 and 2000 to conclude that human resistance has increased from 8% before introduction of the animal drug use to 12% after. [Likewise, if the human rate had been 0 prior to 1990, it would be a fallacy (the *post hoc* fallacy) to assume that the higher rate in 2000 is *caused by* animal drug use, rather than by human drug use or

perhaps by some other cause.] Similarly, if resistance rates increase between two years in some areas while decreasing in others, then referring to an overall "trend" for the whole country based on the pooled data from those areas is misleading, in that no such trend exists at the disaggregate level. These and other logical fallacies in causal inference (e.g., the "Texas sharp shooter fallacy", in which random clusters are pointed to after data have been collected and interpreted as significant) are well recognized (www.fallacyfiles.org/noncause.html) and easy to avoid. Yet, critical examination of recent literature (e.g., FDA, 2001; Angulo *et al.*, 2004; Iovine and Blaser, 2004; Molbak, 2004) suggests that they are surprisingly prevalent in antimicrobial risk assessments and risk management discussions and are sometimes even offered as justifications for recommended risk management policies.

- *Literature citations.* Most hazard identification sections of risk assessment reports include reviews of relevant literature. In preparing these sections, it is important to bear in mind that (a) It is often possible to find peer-reviewed literature to support almost any position or conclusion about a hypothesized health hazard; and (b) The raw data used in a study may imply conclusions different from those emphasized by the study authors in writing up results for publication. Therefore, in addition to summarizing the range of views reported, with their supporting rationales, it is usually important to critically review the rationales and to discuss *why* conclusions differ (e.g., due to differences in modeling methods, operational definitions and laboratory methods for ascertaining endpoints, criteria for causality in stating interpretations and results, etc.). A thorough review should seek to reconcile conflicting results where practicable, e.g., via independent review of raw data. Simply quoting selected excerpts from articles to support particular positions about hazards may be more misleading than informative, and should not be used as an approach to hazard identification. (Against this advice, some risk assessors feel that they should simply report what scientists have said, without attempting to engage in independent analysis and formation of conclusions. However, since risk assessment is peculiarly concerned with creating valid causal models of probabilistic exposure-response relations to help improve risk management decisions, risk assessors may have a unique role to play in drawing on available scientific information and using it to create and validate causal models.)

Example: Limitations of Literature Quotes

Effler *et al.* (2001) stated that "In matched logistic regression analyses, eating chicken prepared by a commercial food establishment in the 7 days before case

illness onset (adjusted odds ratio [AOR], 1.8; P = .03) and consuming antibiotics during the 28 days before illness onset (AOR, 3.3; P = .03) were significant independent predictors of illness." However, these authors did not emphasize another striking result: that eating chicken prepared at home was associated with a 40% reduction in campylobacteriosis risk (relative risk = 0.6) (*ibid*, Table 1.) Thus, the results that are emphasized are not necessarily all of the significant results.

As a second example, a paper by Tollefson and Karp (2004) begins with the assertion that "There is accumulating evidence that the use of antimicrobials in food-producing animals has adverse human health consequences." This regulatory perspective contrasts with conclusions from some other recent publications, such as that "Even though antibiotics have been fed for nearly 50 years to literally billions of animals, there is still no convincing evidence of unfavorable health effects in humans that can be directly linked to the feeding of subtherapeutic levels of antibiotics to swine or other animals" (Cromwell, 2002). It is useful for risk assessors to understand and explain the sources of such disagreements. The literature review sections of hazard identifications for specific animal antibiotics can do so in the context of available facts and data for specific drugs. This topic is explored further in the case study below. [In brief, we believe that the Tollefson-Karp position, which is echoed by Angulo *et al.* (2004) and by WHO (2003), is arrived at by treating unvalidated models and assumptions as if they were data, i.e., the "accumulating evidence" referred to consists largely of expert judgments, speculations, interpretations, and opinions of anti-animal drug use advocates. By contrast, the "convincing evidence" referred to by Cromwell appears to refer more specifically to empirical evidence, i.e., facts and data rather than expert opinion and judgment or speculation. This difference in the types of evidence relied on helps to explain the contrasting conclusions.]

- *Dossier approach.* Presenting information about a drug, bacterium, or other suspected hazard is sometimes called "hazard identification" or "hazard characterization", although "hazard description" might be more accurate. Describing a hazard does not establish the causal link between exposures to it and resulting adverse health consequences that is the essence of hazard identification for applied risk assessment. Presenting detailed information about a drug (e.g., its chemical name, mode of action, pharmacokinetics and pharmacodynamics, and mechanisms of resistance in relevant bacteria) also does not constitute hazard identification, although it is sometimes misleadingly referred to as such.

Example: Information Compilation *vs.* Hazard Identification

Many regulatory agencies request information elements such as those in the following example table for antimicrobial risk assessments.

Table 3: Examples of Requested Information Elements

Information Element	Example
PRODUCT DESCRIPTION	
Chemical name and synonyms	Tylan soluble, tylosin tartrate, (2R, 3R)-2,3-dihidroxy-butanedioate (salt), CAS number: 1405-54-5; Tylan premix, tylosin phosphate, CAS number: 1405-53-4
Class of drug	Macrolide. (16-membered-ring) (Avcare, 2003)
MODES OF ACTION	Protein synthesis inhibitor. Some macrolides may also increase host defenses and interfere with bacterial adherence, enzyme production, and motility, even below the MIC. (Avcare, 2003)
Type of action	"The action of macrolides is usually bacteriostatic, but at elevated doses they can be bactericidal, particularly in the most susceptible species of bacteria." (Avcare, 2003)
Spectrum of activity and MICs	See Avcare, 2003
PHARMACOKINETICS	See http://www.elanco.com/us/pdf/mico300.pdf
PHARMACODYNAMICS	bacteriostatic except at high concentrations
USE IN CHICKENS	
Route and dosing regimen	Feed premix or water
Withdrawal period	1 day for tylan soluble; none for tylan premix (http://www.elanco.com/us/species_poultry_tylan_sol.jsp)
Extent of use	About 9% of broilers
RESISTANCE POTENTIAL	Low. Macrolide resistance rates in people are relatively low and stable in most countries.
Potential mechanisms of resistance in *C. jejuni* and *C. coli*	Mutation of 23S rDNA; CmeB multidrug efflux pump Macrolide-lincosamide-streptogramin (MLS) resistance
Transfer of resistance from chickens to people	Low for *C. coli* and *C. jejuni* (Willems *et al.*, 2000).
Cross resistance to other drugs	Streptogramin and MLS cross resistance
Baseline prevalence of resistance in human patients	About 1%

However, the requested information does not necessarily play any role in the rest of the risk assessment process. For example, FDA-CVM (2003) presents a default regulatory process in which the final risk assessment rating may be entirely driven by how "important" (in some sense that has no clear relation to these requested information elements) the antibiotic class in question (such as macrolides) is judged to be in human medicine. In such systems, the "hazard characterization" information collection process is in danger of becoming little more than an empty ceremony, in which information is collected but not used, and the risk assessment process can be said to have completed this step even though it plays no clear role in identifying interventions that are likely to cause desirable health outcomes.

- *Attributable risk calculations.* A population attributable fraction (PAF) may indicate a positive exposure-effect association, even if there is no causal relation between exposure and effect (or even when the relation is

negative, as in Simpson's Paradox.) For example, if men are more likely than women to eat chicken in restaurants, and if men are also at higher risk than women of campylobacteriosis (but men who eat chicken are at lower risk than men who do not), then the PAF for the association between eating chicken in restaurants and risk of campylobacteriosis could be significantly positive, even if the causal impact of eating chicken is to significantly reduce risk. Unfortunately, it is still common practice in epidemiology to misinterpret PAFs as the fraction of cases that would be prevented if exposure were removed. This interpretation is incorrect in general, but is very widespread, even in otherwise reputable textbooks. The practical implication is that traditional attributable risk and attributable fraction calculations should not be used for purposes of hazard identification in risk assessment.

- *Opinion- and speculation-based hazard identification.* Occasionally, researchers present speculations as if they were a factual basis for hazard identification, or assert that exposure X "may cause" or "can cause" adverse effect Y when it is only not known that X does not cause Y, or when it is only known that X might cause Y under extreme or unusual conditions. Such statements are potentially misleading to decision-makers and should be avoided by clearly stating when hazards are only conjectured or depend on unusual conditions. Good examples of such clear, appropriately caveated, statements can be found in the FDA-CVM, 2004 draft risk assessment for virginiamycin. Even though current empirical evidence does not establish that there is a health risk to humans from use of virginiamycin in animals, it is worth asking how large the risk might be *if* it exists. Such contingent risk assessments allow hazard identification to be bypassed by simply assuming that a hazard exists. This can be useful as long as the underlying assumptions are clearly stated and resulting risks are presented as being contingent on the assumptions. They should not be presented as if underlying assumptions were facts, as occurred in the following case study.

7.1 A Case Study: Presenting Real vs. Hypothesized Hazards

The following case study shows how risk estimates that violate the preceding principles can potentially mislead. Barza and Travers (2002) reported that "Calculations based on estimates of the annual infection rates and attributable fractions of infections [with] *Campylobacter jejuni* suggest that resistance to antimicrobial agents results annually in … an additional 17,668 *C. jejuni* infections, leading to 95 hospitalizations." This statement suggests a causal relation ("results in") between resistance and additional adverse health impacts. It was cited by WHO (2003) (see also Angulo *et al.*,

2004) as quantitative evidence of human health harm from resistance to antimicrobial agents as part of a rationale for recommendations to restrict continued use of animal antibiotics. Angulo *et al.*, 2004 subsequently wrote that "Evidence is also accumulating that the anti-microbial resistance among bacteria isolated from humans could be the result of using anti-microbial agents in food animals and is leading to human health consequences. These human health consequences include: (i) infections that would not have otherwise occurred and (ii) increased frequency of treatment failures and increased severity of infection. Increased severity of infection includes longer duration of illness, increased frequency of bloodstream infections, increased hospitalization and increased mortality." This position has been echoed by regulators advocating the termination of many current animal antibiotic uses (e.g., Tollefson and Karp, 2004).

But the supporting calculations behind the quantitative impact estimates provided by Barza and Travers are based on the following crucial judgments and assumptions, presented here with critical commentary:

1. The calculation began with the following statistics: "There are 2,453,926 infections, 13,174 hospitalizations, and 124 deaths caused by *C. jejuni* each year in the United States (Mead *et al.*, 1999)." Even before examining these specific numbers, it is clear that the term "caused by" is being used very loosely here. *Any* hospitalization or death that occurs in a patient with campylobacteriosis is assumed to be *caused* by campylobacteriosis. As explained by Mead et al., "To estimate the number of deaths *due to* bacterial pathogens, we used the same approach described for hospitalizations: first calculating the number of deaths *among* reported cases, then doubling this figure to account for unreported deaths, and finally multiplying by the percentage of infections attributable to foodborne transmission" (emphases added. Note the conflation of "due to" and "among".) Since having campylobacteriosis does not confer immortality, some patients with campyobacteriosis will die, but this should not be taken to imply that campylobacteriosis caused their deaths. To the contrary, severe illnesses such as leukemia and AIDS that compromise the immune system are strong risk factors for *Campylobacter* infections, as well as for hospitalization, other infections, and early deaths (Sorvillo *et al.*, 1991, Cox, 2003). Thus, attributing hospitalizations and deaths among such severely ill patients to *Campylobacter* rather than to the underlying severe illness has the effect of blaming some AIDS-related deaths and deaths from other severe illnesses on campylobacteriosis infections that happen to be present but that played no role in causing death. The starting numbers are thus inflated by an unknown amount.

2. Turning now to the specific numbers cited, the total estimated total infections reported by Mead et al. were extrapolated from a source that they cited as follows: "Passive surveillance estimate based on average number of cases reported to CDC, 1992-1994 (CDC, unpub data). Active surveillance estimate based on extrapolation of average 1996-1997 FoodNet rate (24.1 cases per 100,000 population) to 1997 U.S. population." Total infections were "assumed to be 38 times the number of reported cases, based on studies of salmonellosis." The rationale offered was that:

   > "For *Salmonella*, a pathogen that typically causes nonbloody diarrhea, the degree of underreporting has been estimated at ~38 fold (Voetsch, manuscript in preparation). …Because similar information is not available for most other pathogens, we used a factor of 38 for pathogens that cause primarily nonbloody diarrhea (e.g., *Salmonella, Campylobacter*) and 20 for pathogens that cause bloody diarrhea (e.g., *E. coli* O157:H7, *Shigella*). For pathogens that typically cause severe illness (i.e., *Clostridium botulinum, Listeria monocytogenes*), we arbitrarily used a far lower multiplier of 2, on the assumption that most cases come to medical attention."

Thus, while Barza and Travers' claim that "There are 2,453,926 infections, 13,174 hospitalizations, and 124 deaths caused by *C. jejuni* each year in the United States (Mead *et al.*, 1999)" appears to provide useful data to policy makers, tracing back the Mead *et al.*, 1999 citation shows that these numbers are only guesses (based largely on Voetsch's inflation factor of 38 in an unpublished, incomplete manuscript for *Salmonella*, applied to unpublished data for *Campylobacter* from a decade earlier.) When WHO (2003) and subsequent authors then cited the Barza and Travers paper as a source of "evidence" on quantitative health risks, the detail that the numbers originated in unpublished guesswork was lost and the distinction between real ("hard") data and speculation started to become blurred.

3. Campylobacteriosis rates in the US have fallen since the unpublished data relied on by Mead *et al.*, 1999 were collected. They may have been closer to 13.4 per 100,000 in 2002 (CDC, 2003) than the 24.1 per 100,000 assumed by Mead et al., and hence by Barza and Travers. This suggests a need to update the starting estimate of "2,453,926 infections, 13,174 hospitalizations, and 124 deaths caused by *C. jejuni* each year in the United States", even if the use of "caused by *C. jejuni*" to mean "with *C. jejuni* present" is not addressed.
4. Barza and Travers continued their calculation as follows: "Eighteen percent of these, or 441,707 infections, are caused by strains resistant to at least one antimicrobial agent." Again, "caused by" is used

imprecisely here. It implies that if a campylobacteriosis case includes even one resistant strain for at least one antibiotic (as determined by an *in vitro* test), then it can be considered to be "the cause" of the illness, regardless of how many susceptible strains are present.

5. "Again assuming that rates of hospitalizations and death are similar for infections with drug-resistant strains and with drug-susceptible strains, then 2371 hospitalizations and 22 deaths result each year from infection with *Campylobacter* strains resistant to at least one antimicrobial agent." This does not indicate that resistance has any bearing on clinical outcomes, making attribution of outcomes to resistance questionable.
6. Perhaps the most important step in the Barza and Travers calculation came next, as follows: "We will assume that the attributable fraction was 5% …*If the attributable fraction is 5%,* this translates to 22,085 infections, 119 hospitalizations, and 1 death in the United States each year as a result of infection by quinolone-resistant *C. jejuni"* (emphasis added). The assumption made here that "the attributable fraction is 5%" is just an *assumption*. No empirical justification is offered for it. It is the modeler's judgment, not an empirical finding. Yet, the whole risk assessment is contingent on this assumption – a fact recognized when the assumption is introduced (as italicized above), but lost when the conclusions are presented (and subsequently cited by WHO). Had Barza and Travers posited 0%, rather than 5%, as the attributable fraction, then this "evidence" of human health harm, subsequently cited by WHO (2003) in calling for reduced use of animal antibiotics, would disappear. In other words, the crucial step in creating quantitative risk estimates in this analysis is not data at all, but a decision by Barza and Travers to assume an attributable fraction of 5%. This example illustrates how personal subjective judgments can produce dramatic-looking numbers that are then cited and used as if they were data, potentially misleading risk managers. To avoid misleading decision-makers, it is essential to label speculations as such and not to conflate them with data.
7. The attributable fraction that Barza and Travers refer to measures *association,* not *causation*. Thus, use of the term "as a result of" in the above quote (i.e., a causal interpretation) is not justified. (As a counterexample, if people who are at higher risk of campylobacteriosis are also more likely to take antibiotics, then the attributable fraction of campylobacteriosis cases associated with antibiotic use may be large even if one does not cause the other.) But Barza and Travers misinterpret this attributable fraction as causal, stating that "The attributable fraction reflects the proportion of all cases that would not have occurred in the absence of recent or concurrent treatment with an antimicrobial agent to which the bacterium was resistant." This

preventable fraction is logically distinct from the *attributable fraction*; in theory they may even have opposite signs.

8. Barza and Travers continue: *"*If 80% of *C. jejuni i*nfections arise from food animals…then antimicrobial resistance in these animals contributes to 17,668 infections and 95 hospitalizations per year." The 80% number can also be traced to guesswork documented in Mead *et al.*, 1999. It comes from a 1992 study indicating that "Although waterborne outbreaks occur, foodborne transmission accounts for most of the sporadic cases". "Most of" was subjectively quantified as "80%". But, improved understanding of *Campylobacter* epidemiology since 1992 suggests that sporadic cases now have many sources, rather than one or a few (especially chicken) that were thought in 1992 to be the dominant contributors. For example, Sopwith *et al.* (2003) concluded that "sources and vehicles of human *Campylobacter* infection are numerous and interventions that target a single risk factor are unlikely to impact significantly on the overall burden of disease." Moreover, Stern and Robach (2003) found that *Campylobacter* loads in processed chicken carcasses had declined by more than 90% since the mid nineties. Thus, the use of 1992 data in Mead *et al.*, 1999 is of doubtful relevance for quantifying risks in 2002, the year of the Barza and Travers paper. At a minimum, the 80% number should be updated to reflect developments since 1992.

This case study, reviewing the detailed derivation of numerical risk estimates in Barza and Travers (2002) and subsequently cited by WHO (2003), indicating that "Resistance to antimicrobial agents results annually in an additional 17,668 *C. jejuni* infections, leading to 95 hospitalizations" shows that these numbers are based on a subjective judgment to "assume that the attributable fraction was 5%", combined with a misinterpretation of *attributable fraction* as *preventable fraction*, i.e., "the proportion of all cases that would not have occurred in the absence of recent or concurrent treatment with an antimicrobial agent to which the bacterium was resistant". What is missing that is essential for sound quantitative risk assessment is empirical evidence that, if antimicrobial resistance were not present, there would be a change for the better in clinical outcomes. (Barza and Travers appear to assume that if resistance were eliminated, illnesses would be reduced by the elimination of currently resistant cases. But it is realistic to assume that eliminating resistance would only replace resistant cases with susceptible ones; thus, the human health effects of these additional susceptible cases must still be considered.) Moreover, the analysis does not quantify the effects of recommended reductions in animal antibiotic use on the rates of susceptible and resistant infections in the human population.

It is impracticable for most consumers of risk assessment information – including policy makers reading the WHO (2003) report and similar documents – to carefully trace all of the references cited and to determine where opinions and speculations (whether by opponents of animal antibiotic use or other interest groups) have been substituted for data. Moreover, even fictitious risk numbers, once cited, tend to acquire a life of their own and to be cited or referred to in further documents as "evidence" providing support for policy positions (e.g., Angulo *et al.*, 2004, Karp and Tollefson, 2004). Therefore, it is incumbent upon responsible risk assessors to carefully distinguish for their readers between facts and assumptions (or opinions, speculations, envisioned possibilities, etc.) in presenting and summarizing their results. Sensitivity analyses, discussed in Chapter 5, can help to show how sensitive risk assessment conclusions are to input assumptions, but it is the responsibility of risk assessors to make clear where assumptions have been used. Otherwise, it is possible for national and international perceptions and risk management policies to be shaped by the "data" and "evidence" offered by stakeholders and cited by regulators without the involved policy makers even realizing that key aspects of the "data" are simply made up.

8. CONCLUSIONS

This chapter has suggested several do's and don'ts for the conduct of hazard identification – the first, and often the most important and most difficult, step of the risk assessment process. Hazard identification is crucial because it deals directly with the key questions of whether a real threat to human health from a given source exists and, if so, whether adverse effects can be reduced by reducing or preventing exposures. It is challenging because it requires drawing (and validating) causal inferences from available data. Once hazard identification shows that a risk is real and preventable, the rest of risk assessment can help to quantify the extent to which health outcomes can be improved by reducing exposures. But the bare hazard identification knowledge that a threat exists and is preventable is often the most powerful information that risk assessment provides to risk managers and policy makers.

The "do's" in the chapter largely focus on recommendations to use appropriate current statistical methods and algorithms for causal analysis to avoid common pitfalls and fallacies in causal reasoning and to help draw sound, trustworthy conclusions about hazard identification. The "don'ts" focus largely on avoiding substituting intuition, speculation, and judgment for rigorous causal analysis of data in arriving at hazard identification conclusions. An important theme is that it is not enough to identify an

apparent non-random association and then use judgmental criteria to decide whether it is likely to be causal. Even though this approach is sometimes explicitly advocated (e.g., Surgeon General, 2004), the available evidence is that human judgment-based approaches often do not work well (since human judgments about causation are notoriously fallible), while even simple quantitative empirical methods often work better (e.g., Plous, 1993). A number of examples throughout the chapter have illustrated how these general principles for hazard identification apply to microbial and antimicrobial risk assessment.

Chapter 4

Exposure Assessment

1. INTRODUCTION TO RISK QUANTIFICATION

Hazard identification critically examines the empirical evidence for whether changing acts really will cause the changes in health consequences predicted by the risk model template in Figure 1 of Chapter 3. Risk quantification, discussed in this chapter and Chapter 5, provides additional information needed for effective risk management by predicting by *how much* different risk management interventions, acts, or decisions will change probable health outcomes. With this information, a risk manager can begin making trade-offs and optimizing among risk management alternatives.

Risk quantification consists of the three steps of exposure assessment, dose-response modeling, and risk characterization (including uncertainty and sensitivity analyses). These are the main topics of this chapter and Chapter 5. In addition, quantifying the threat of resistance to antibiotics requires considering the dynamics of the system composed of animals (ill and well), humans (ill and well), and bacteria (both resistant and susceptible) that flow among animal and human hosts. This chapter therefore further applies the simple systems dynamics model of foodborne bacterial illnesses and resistance introduced in Chapter 3 and places it within the framework of more traditional microbial exposure assessment modeling, in which the microbial load reaching consumers via food servings is expressed as a product of factors that can be estimated from data.

2. INTRODUCTION TO EXPOSURE ASSESSMENT

Exposure assessment has been described as "the qualitative and/or quantitative evaluation of the degree of intake likely to occur" (WHO/FAO, http://www.who.int/foodsafety/micro/riskassessment/en/). Although qualitative risk assessment has the problems noted in Chapter 1, qualitative assessment of exposures is often more useful and less problematic, as discussed at the end of this chapter. In the causal chain

$$\text{act} \rightarrow \Delta\text{exposure} \rightarrow \Delta\text{illnesses} \leftarrow \text{covariates},$$

exposure assessment describes the "act $\rightarrow$ Δexposure" link. It estimates population exposures to microbial loads for different risk management acts. Individual exposures depend on the numbers of colony-forming units (CFUs) of different bacteria (resistant and susceptible) ingested per unit time via food, water, from contaminated hands, and via other pathways. For populations, exposure refers to the frequency distribution of individual exposures (microbial loads) consumed per unit time.

2.1 Definition and Purposes of Exposure Assessment

The US FDA has defined exposure assessment as "A component of a risk assessment that characterizes the source and magnitude of human exposure to the pathogen". The magnitude of human exposure, also called the dose, is defined as "The amount or number of a pathogen that is ingested or interacts with an organism (host)". For additional discussion, see: http://www.foodsafety.gov/~dms/lmriskgl.html. These are roughly analogous to concepts used in environmental risk assessment. For example, US EPA experts have stated that "Questions raised in the exposure analysis concern the likely sources of the pollutant… its concentration at the source, its pathways (air, water, food) from the source to target populations, and actual levels impacting target organisms" (Patton, 1993).

Exposure assessment has the following goals:

- Identify exposed subpopulations at risk of infection and illness from exposures to hazards
- Identify conditions leading to high-risk exposures
- Describe the extent of exposures (frequency and magnitude of individual exposure in the population in relation to susceptibility and covariates)
- Predict how risk management decision options will affect exposures.

A successful exposure assessment therefore describes the frequency distribution of microbial loads ingested by members of exposed populations

and subpopulations. It should show how these distributions change for different risk management decisions. The descriptions should contain enough detail to discriminate among different microbial load distributions that would cause significantly different health outcomes. This information is used, together with dose-response information, in risk characterization (see Chapter 5.)

A useful surrogate for exposure in many cases is the *number of servings* containing potentially infectious doses of the bacterium of concern ingested per year (for population exposure) or per capita-year (for individual exposure). A "potentially infectious dose" is any dose large enough to infect a susceptible consumer. It may be as small as one CFU, if that is biologically realistic. If reliable dose-response information shows that the risk of illness below some number of ingested CFUs is negligible (or is small enough so that it can be ignored without changing the expected number of illnesses per year, within the limits of rounding error), then a "practical threshold" – meaning one that leads to numerically accurate risk calculations – may be used, even if, in principle, no true biological threshold exists. The number of servings per year ingested with microbial loads above the practical threshold then defines total annual exposure.

Example: Calculating Risk with Surrogate Exposure Variables

Setting: Suppose that the frequency distribution of microbial loads of a pathogen (e.g., *Campylobacter* or *Salmonella*) ingested in servings of a food commodity (e.g., chicken, eggs, or salads) is uncertain, but that available information constrains the distribution as follows:

- In at least 95% of servings, the load is small enough to create a minimal illness risk of not more than 1×10^{-6} (i.e., most servings are relatively low-risk)
- In at most 1% of servings, load is large enough to create an illness risk of 1% or more (i.e., at most 1% of servings are relatively high-risk)
- For other servings, load is intermediate between the preceding two cases.

Problem Calculate a plausible upper bound on the risk (measured as expected number of illnesses per serving) caused by this uncertain exposure profile.

Solution: An upper bound on risk is obtained by making a worst-case bounding assumption that *all* exposures that do not create "small" risks (less than 1×10^{-6}) create the maximum possible risk of 100% probability of illness. Since at most 5% (i.e., $1/20^{th}$) of servings fall in this category, the upper bound on risk is 1 expected illness for every 20 servings. This calculation illustrates the use of simple bounding assumptions to obtain a conservative risk estimate for servings as a surrogate exposure variable, even though actual microbial loads ingested remain unknown.

3. EXPOSURE ASSESSMENT METHODS: SIMULATION

The approximate amount of food-borne bacteria ingested per day, per serving, or per meal, and the approximate shape of the frequency distribution of microbial loads ingested, relative to the dose-response relation (e.g., how frequent are exposures that are likely to cause illness?) drives quantitative risk. This reflects the precept that, even for bacteria, "the dose makes the poison". However, in practice, it is quite common for the microbial loads received to be very uncertain, especially if they depend on unmeasured and/or highly variable processes such as cooking of food, cross-contamination of other foods in the kitchen, or transfer from contaminated surfaces to skin to ingestion. In such cases, the exposure assessment causal diagram may look like this:

$$\text{act} \rightarrow \text{exposures} \rightarrow \text{illnesses} \leftarrow \text{individual covariates}$$
$$\downarrow$$
$$\text{measured exposure surrogates}$$

Available data consist of surrogate measurements (e.g., microbial concentrations in carcass rinses at retail, or on swabbed surfaces) rather than direct measurements of ingested microbial loads

Exposure modeling for applied risk assessment with surrogate exposure measurements consists of estimating how the underlying true exposures will change if different risk management actions are taken, while subsequent exposure-response modeling must focus on how health risks will change when true exposures are changed by decisions. True exposures then play the role of *latent variables* in causal modeling, i.e., variables that affect observed outcomes but that are not directly observed themselves. Appropriate statistical techniques for causal diagrams with latent variables (e.g., Shipley, 2000 for linear models; Pearl, 2002 and Hartemink *et al.*, 2001 for more general Bayesian Network models) can be applied to the above diagram with surrogate measurements of exposure for data. Software such as WinBUGS helps to automate the required computations for inference with missing data and unobserved or surrogate variables.

Statistical issues aside, exposure assessment uses *predictive microbiology models* to predict how microbial loads reaching consumers or other exposed populations (e.g., patients) change if different risk management decisions are made. Such a model predicts how microbial loads change with time due to cell divisions, survival and death as a function of conditions such as temperature, moisture, growth medium (e.g., type of food in which the bacteria are growing), and physical and chemical environment (e.g., pH) for bacteria growing in processed or stored foods.

Example: Predictive Microbiology Data and Models

Table 1 shows the estimated decrease in microbial load of different strains of *Campylobacter* for different storage temperatures and food substrates. Similar tables, as well as mathematical models and figures based on them, describe standard bacterial growth curves under different conditions (such as pH, salinity, and ambient temperature) for different foods and bacteria, as well kill curves (e.g., fraction of bacteria killed *vs.* cooking times) for different cooking temperatures and conditions (e.g., Haas *et al.*, 1999; van Gerwen *et al.*, 2000; Oscar, 2004). From such data, one can create statistical models of the probable number of bacteria on meat after different initial contamination, storage, processing and cooking histories. Simulating different histories then creates corresponding distributions of predicted resulting exposures (i.e., microbial loads in ingested servings at the end of these processes.) However, in reality, it is seldom practical to collect all the data needed to accurately simulate the distribution of contamination, storage, and preparation histories, and simpler approaches are used, as described next.

Table 1: Effect of chilling and freezing on *Campylobacter* in meat products

Substrate	Storage temp (°C)	Initial decrease (log_{10} cfu/day)	Total decrease (log_{10} cfu/day)	Strains examined
Chicken carcass	-20	-0.1-1.4/21	-0.5-2.3/84	5C.j./C.c.
Chicken drip	-20	-01-1.1/21	-0.6-2.5/84	5C.j./C.c.
Chicken carcass	-20	-0.5/36	-1.4/64	NF
Chicken liver	-20	-1/'few'	-1.6/85	NF
Chicken drumsticks	-20	-1.4/7	-2.7/182	1 C.j.
Chicken breast skin	-20	-2.4/3*	ca.–3.7/56	1 C.j.
Ground beef liver	-20	-0.9-1.4/3	-2.3-2.6/84	5 C.j./C.c.
Ground beef	-15	-3/3	-3/14	5 C.j.
Raw chicken breast	2	-	-5-6/24	2 C.j.
Raw minced beef	2	-	-5-6/27	2 C.j.
Cooked minced beef	2	-	-5-6/49	2 C.j.
Patê	2	-	-5-6/15	2 C.j.
Ground beef liver	4	-0.0-0.4/6	-	5C.j./C.c.
Cooked chicken	4	-0.3-0.7/7*	-	3 C.j.
Chicken carcass	4	-0.6-1/4-7	-	NF
Chicken drumsticks	4	-0.7/7	-	1 C.j.
Chicken breast skin	4	-1.4/7*	-	1 C.j.
Raw chicken breast	10	-	-5-6/13	2 C.j.
Cooked minced beef	10	-	-5-6/23	2 C.j.
Patê	10	-	-5-6/6	2 C.j.

C.j. = Campylobacter jejuni; C.c. = *Campylobacter* coli; *, numbers estimated for the reference; NF, natural *Campylobacter* contamination

Source: Christensen *et al.* (2001), p. 11

3.1 Practical Simulation-Based Exposure Modeling

Exposure models describe the transport, growth, and spread of hazardous materials, such as concentrations of pathogenic bacteria on food animal products, through different media and pathways (e.g., foods, drinking water) leading from their source(s) to members of the exposed population. In addition, exposure models may consider the distribution over time of human populations among locations (e.g., restaurants and kitchens) and activities (such as purchasing, handling, and preparing foods) that result in exposures.

Exposure assessment data and calculations can be organized and presented using any of the following major exposure-modeling approaches.

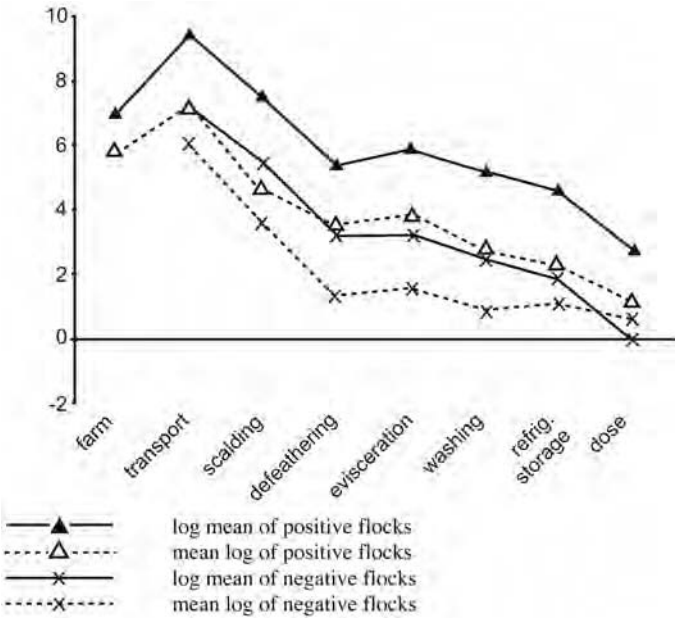
1. Product of factors approach. The food production process is divided into a sequence of points at which microbial loads can be measured or estimated, such as: in chickens leaving the farm, chickens arriving at slaughter following transportation, chicken carcasses exiting the slaughter plant, chicken carcasses or servings at the point of retail, chicken carcasses or servings entering the kitchen, and chicken servings at the point of ingestion. Then, the increases or decreases in microbial loads between consecutive points are modeled by multipliers that can be estimated either empirically (from the ratios of the loads at each point compared to its predecessor) or by a predictive microbiological model. The product of these multipliers predicts the microbial load at the last point (ingestion) from the microbial load entering the first point. Often, the multipliers are treated as random variables (e.g., with approximately log-normal distributions), and the fraction of microbial loads at ingestion that exceed a desired level (e.g., a characteristic infectious dose level) can then be assessed using probability theory or by Monte Carlo simulation using commercial risk analysis software such as Analytica™, Crystal Ball™, or @RISK™ (Vose, 2000; Clemen and Reilly, 2000).

Example: Product-of-Factors Approach for *Campylobacter* Exposures

Figure 1 shows model-based estimated change in the logarithm of the numbers of *Campylobacter* cells per chicken carcass during processing of fresh air-chilled carcasses, from both *Campylobacter*-positive and *Campylobacter*-negative flocks. As noted by the authors: “The model indicated that the external *Campylobacter* load per chicken increased during transport and evisceration, and decreased at the other processing steps studied, with an overall reduction of the mean load from farm to fork of about 4 to 5 logs [Figure 1]. The prevalence of *Campylobacter*-contaminated chickens from positive flocks appears to drop from 100% of live birds to 20% of chicken meat servings… . For negative flocks, prevalence increases during transport, defeathering and evisceration, indicating the

effect of cross-contamination during processing. Prevalence later drops to a value of about 3% of servings at the moment of consumption."

Figure 1: Estimated Log_{10} of *Campylobacter* Loads at Successive Stages



Negative flocks get contaminated during transport. Both the mean of the logs and the log of the means are given. (The differences between these is as a result of the skewness of the distribution of values and the fact that 'zero'- values cannot be incorporated in calculations of the mean of logs (which is therefore only about the positive carcasses)). *Source* FAO/WHO, 2002. ftp://ftp.fao.org/es/esn/food/cv_02e.pdf

2. Process simulation modeling approaches (Haas *et al.*, 1999, 225-248) describe the flow of food animal carcasses, portions, and products through various sub-processes, each characterized by an input-output relation. These component input-output relations may be described by simple empirical regression models or by other statistical or simulation models. Changes in microbial loads from step to step of the process are tracked (e.g., via discrete-event simulation, or by sampling from the probability distribution for the multiplier by which loads are increased or decreased at each step.) Available measurements and data may be used to fit simple probability distributions and parametric models to describe the growth or attenuation of transmitted microbial load at each stage. Examples include Poisson or negative binomial distributions of microbial loads (fit using most probable number (MPN) data and maximum-likelihood statistical estimation algorithms) and Gompertz growth curves for pathogen growth kinetics. Data

for growth rates of *E. coli* O157:H7, *Listeria monocytogenes, Clostridium botulinum, Staphylococcus aureus,* and other common pathogens are found in Haas *et al.* (1999, Chapter 6) and its references. Consumption factors and frequencies for water and foods (beef, fish, chickens, eggs, shellfish, etc.) are available from the literature (*ibid,* p. 239, 241) and can be used to model the frequency with which microbial loads on food portions are ingested. Transfer rates of bacteria between skin and hands and from food to hands have also been estimated for various bacteria; however such details are not necessarily needed or useful if adequate data on earlier and later points in the causal chain leading from animal loads to human illness are available.

3. Farm-to-fork models are an important type of process simulation model. Farm-to-fork models track the microbial load distributions on animal carcasses, portions, and servings through successive stages of production, processing, transport, slaughter storage, preparation, and consumption. Monte Carlo simulations of the probabilistic input-output relations at each stage are used to propagate microbial load frequency distributions throughout the model. A closely related methodology is *dynamic flow tree modeling* for microbial risk assessment (Marks *et al.*, 1998). Dynamic flow trees use Monte Carlo analysis to sum risks over many scenarios, weighted by their respective expected frequencies or probabilities, but without necessarily representing all the process steps in a farm-to-form model.

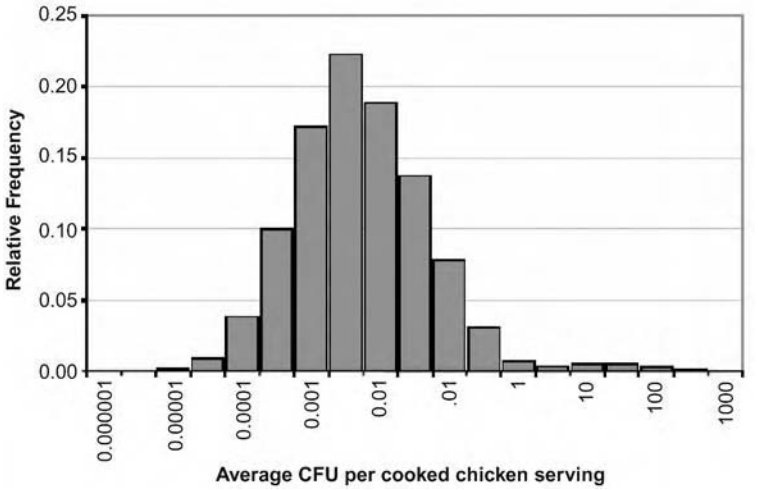
Example: Farm to Fork Exposure Models for Pathogens in Chickens

Fully worked-out examples illustrating the farm-to-fork approach can be found in the *Campylobacter* risk assessment of Christensen et al, 2001 and in a 2002 *Salmonella* risk assessment (starting at slaughter rather than at the farm) provided by WHO/FAO (http://www.who.int/foodsafety/publications/micro/Salmonella/en/) The latter summarized its exposure assessment sub-model as follows:

> "The exposure assessment of *Salmonella* in broiler chickens mimics the movement of *Salmonella*-contaminated chickens through the food chain, commencing at the point of completion of the slaughter process. For each iteration of the model, a chicken carcass was randomly allocated an infection status and those carcasses identified as contaminated were randomly assigned a number of *Salmonella* organisms. From this point until consumption, changes in the size of the *Salmonella* population on each contaminated chicken were modeled using equations for growth and death. The growth of *Salmonella* was predicted using random inputs for storage time at retail stores, transport time, storage time in homes, and the temperatures the carcass was exposed to during each of these periods. Death of *Salmonella* during cooking was predicted using random inputs describing the probability that a carcass was not adequately

cooked, the proportion of *Salmonella* organisms attached to areas of the carcass that were protected from heat, the temperature of exposure of protected bacteria, and the time for which such exposure occurs. The number of *Salmonella* consumed were then derived using a random input defining the weight of chicken meat consumed, and the numbers of *Salmonella* cells in meat as defined from the various growth and death processes. Finally, in the risk characterization, the probability of illness was derived by combining the number of organisms ingested (from the exposure assessment) with information on the dose-response relationship (hazard characterization)."

In the *Salmonella* risk assessment, the results of Monte Carlo simulation exposure modeling are presented as: (a) A prevalence of contaminated broilers at the point of serving (estimated as 2%); and (b) The following conditional frequency distribution for the dose (CFUs)-per-serving from contaminated broilers:



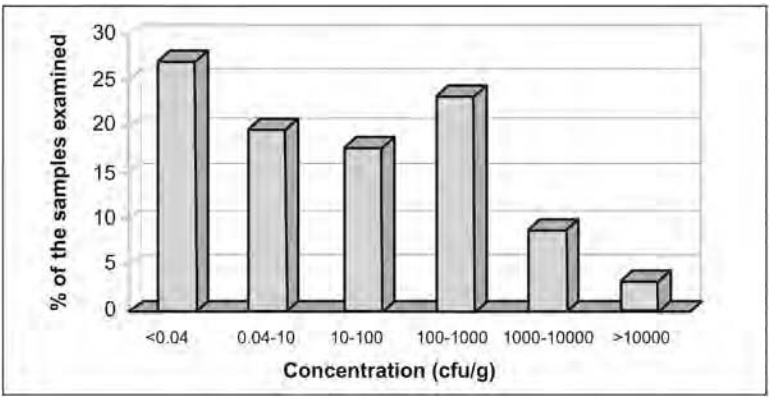
Source: http://www.who.int/foodsafety/publications/micro/*Salmonella*/en/

This frequency distribution, reproduced from the results section of the risk assessment, shows how large an exposure a person is likely to receive from a serving of contaminated, undercooked broiler chicken. This distribution is the main output of the exposure assessment and the main input to the dose-response model for purposes of calculating the illness risk per serving.

Example: Monte Carlo Simulation of Exposure Distributions

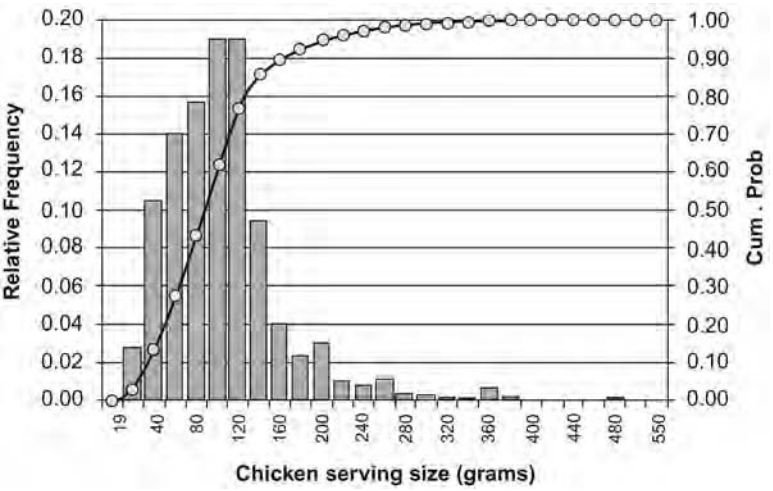
In practice, information on frequency distributions of microbial loads may only be available from samples of retail products, rather than from servings at the point of ingestion. For example, Christensen *et al.* (2001) provide the following

empirical frequency distribution for *Campylobacter* load concentrations (expressed in CFUs per gram) in retail samples:



Source: Christensen *et al.* (2001), p. 11

Starting with this retail load distribution, predictive microbiology models (combined with input data on storage, preparation, and consumption behaviors, including the frequency distribution of serving sizes, in grams) can be used to generate corresponding predicted frequency distributions of microbial loads at ingestion. Specifically, suppose that the frequency distribution of chicken serving sizes is as shown in the figure below (from WHO/FAO, 2002, Figure 6.11, p. 186).



Source: http://www.who.int/foodsafety/publications/micro/Salmonella/en/

Then the frequency distribution for the number of CFUs ingested per serving can be found from these inputs by Monte Carlo simulation. This is done as follows:

1. Randomly sample a *serving size* (grams per serving), drawn from the serving size distribution.

2. Independently, randomly generate a *serving contamination level* (CFU per gram at the point of ingestion), formed by randomly sampling a contamination concentration from the contamination concentration distribution and then multiplying it by an estimated load reduction factor for storage (e.g., chilling, as in Table 1) and cooking, sampled from an empirical distribution of such factors derived from predictive microbiology models or data.
3. Multiply the values of the serving size and the serving contamination level to obtain a sample value for CFUs ingested per serving.
4. Iterate steps 1-3 to obtain a frequency distribution of CFUs ingested per serving.

This Monte Carlo simulation procedure for estimating the distribution of a product from the empirical distributions of its components (or from their conditional distributions, if the values of some affect the likely values of others) can be implemented easily using commercial risk analysis tools (Vose, 2000).

3.2 Simulation Based on Conditioning, Not Imitation

A common source of confusion about quantitative exposure models (and other components of quantitative risk assessment) is a misperception that they require unrealistically detailed data to implement (e.g., Bartholomew et al., 2005). This is not true if appropriate techniques of uncertainty analysis are used to bridge data gaps in the modeling. For example, suppose that there are several consecutive stages in a food safety process simulation exposure model, such as:

$$A \Rightarrow B \Rightarrow C \Rightarrow D$$

Here, the thick arrows represent input-output processes and A, B, C, D are points where microbial loads might be measured – e.g., A = on animals leaving the farm, B = on animals following transportation to the slaughter facility, C = post-processing carcass, D = retail meat product. If measurements are unavailable for stage C, then the conditional probability distribution of microbial loads at D can still be related to microbial loads at A (thus leaving the chain unbroken by the missing data at C) via the conditional probability formula:

$$\Pr(\text{load at D} = d \mid \text{load at A} = a) = \Sigma_b[\Pr(\text{load at D} = d \mid \text{load at B} = b) \times \Pr(\text{load at B} = b \mid \text{load at A} = a)]$$

where

$$\Pr(\text{load at D} = d \mid \text{load at B} = b) = \Sigma_c[\Pr(\text{load at D} = d \mid \text{load at C} = c) \times \Pr(\text{load at C} = c \mid \text{load at B} = b)]$$

In other words, it is possible to *condition on what is observed* while skipping over (or "marginalizing out", in statistical terminology) the unobserved quantities by summing over all their possible values, weighted by their conditional probabilities. This makes predictive modeling possible even when only some of the information needed for a complete description of the process being modeled is available (Richardson and Green, 1997).

This simple example illustrates an important point about the data requirements for risk assessment modeling. The basis for inferring exposure distributions from available data is usually statistical, i.e., *conditioning* of probability distributions for the output quantities of interest in a causal model (the simulation model), on what is known and measured. What is measured may be very imperfect, such as rough exposure estimates or surrogates (e.g., how many times per week individuals eat undercooked servings of meat). Algorithms for performing conditional probability calculations to infer the conditional probability distributions for quantities of interest, conditioned on measured values, in specified causal graph models are now well understood (e.g., Zhang, 1998) and are becoming more widely available in commercial statistical software. The usual basis for inference is *not* exhaustive description and simulation of all relevant physical details of the processes that lead to exposures, which could indeed impose unrealistic data collection requirement in many cases.

Example: Mixture Distributions and Unknown Dose-Response Models

The two main components of a health risk assessment model are the *exposure model*, which predicts the frequency distributions of exposure units ingested in the population per unit time, for each of the different risk management decision options being compared; and a *dose-response model* (or exposure-response model) that predicts the expected number of illnesses per unit of exposure ingested (e.g., per contaminated serving consumed, if illness probability is not sensitive to the amount of contamination above some threshold; or per CFU consumed, if the dose-response model is linear with no threshold.) The dose-response function may depend on many unknown or unobserved factors, such as the exposed individual's immune status and susceptibility to the particular bacterium of concern. The mathematical form of the dose-response relation may also be unknown. Despite these unknowns, risk may be decomposed conceptually as follows:

$$\Pr(\text{Illness} \mid \text{exposure} = x) = \sum_r \Pr(\text{Illness} \mid \text{exposure} = x \ \& \ \text{response type} = r) * \Pr(\text{response type} = r)$$

where "response type" is an unobserved variable summarizing all of the missing information needed to predict the probability of illness from a known level of

exposure. (For example, if each individual has an unknown threshold number of CFUs that must be ingested in one meal to cause illness, then r would be that threshold number. If there is a continuum of response "types", the above sum is replaced by an integral.) An important development in mathematical statistics is the recognition that the uncertain quantities Pr(response type = r) can be interpreted as statistical *coefficients* to be estimated directly from data on the aggregate number of responses observed in populations for different exposure conditions, while the conditional response probabilities that are paired with these coefficients, Pr(illness | exposure = x, type = r) can be estimated simultaneously from the same data (provided that technical identifiability conditions are met. These are automatically satisfied by many large families of distributions.) The required statistical technology is that of *finite mixture distribution models* if the number of types is finite; or continuous mixture models if types are continuous. In the conceptual model $\Pr(\text{Illness} \mid x) = \sum_r \Pr(\text{Illness} \mid x, r) * \Pr(r)$, the coefficients Pr(r) are interpreted as mixing coefficients. Pr(Illness | x, r) is the conditional response probability given latent (unobserved) variable r. Well developed computational Bayesian algorithms can be applied to estimate the number of components in the mixture (i.e., the number of statistically significantly different "types") and the corresponding coefficients and conditional response probabilities (see e.g., Richardson and Green, 1997.) Note that, in this construction, the definition of the exposure variable x can be any measured quantity (e.g., CFUs in processed carcasses or in retail meats) that can be paired with corresponding illness rates. All unobserved details of subsequent processing, handling, preparation, etc. are then absorbed into the latent "type" variable, r. Missing values and errors in measured values of x can also be handled within the computational Bayesian framework (e.g., using the data augmentation algorithm, Schafer, 1997) to allow the conditional distributions of outputs given observed data to be quantified, even when other data are missing. There is thus great flexibility within simulation approaches to use all available data (via conditioning), but without requiring use of unavailable data.

Scientists and modelers not trained in risk assessment sometimes mistakenly assume that *physical-level* simulation of hard-to-model and unobserved processes such as cooking, cross-contamination, and microbial growth and death curves under various conditions, are required to obtain useful, valid models of exposure. They therefore reject simulation-based exposure modeling, feeling that its input requirements are unrealistic and/or require making speculative assumptions about unmeasured processes. Understanding that all that is necessary is to compose valid statistical descriptive relations (i.e., conditional probability relations for measured quantities), as in the above example, may help to alleviate these concerns.

4. ATTRIBUTION-BASED EXPOSURE MODELING

In contrast to the preceding approaches, *retrospective attribution models* begin with empirical estimates of the number of cases per year of adverse health outcomes caused by bacteria of concern. They then use genotyping, serotyping, resistance typing, food consumption rates, and other data to estimate the fractions of these cases of adverse effects that could be prevented by removing specific exposures and sources. As explained in Chapter 3, these *preventable fractions* should not be confused with the population *attributable fractions* (PAFs) often calculated in epidemiology, nor with the fractions of exposures that have their historical origins in various sources. For example, PAFs for multiple causes can sum to more than 100%, since they only reflect association, not causation; but preventable fractions cannot exceed 100% for any intervention. As another example, it is logically possible that use of a certain animal drug might have caused (or helped to cause) a resistant strain of a bacterium to emerge, and yet terminating the drug use now might not reduce the resistant strain if it competes favorably with other strains even in the absence of continued selection pressure. In this case, the fraction of resistance caused by the historical drug use (perhaps as high as 100%) would not help to predict the fraction of resistance that would be prevented by discontinuing further drug use. It is the latter – the preventable fraction of future exposures to the resistant strain – that is of greatest interest in deciding what to do now, and it is this fraction that attribution-based approaches should seek to quantify.

In practice, precise quantification of preventable fractions from existing data may be difficult or impossible, since the effects on illness rates of counterfactual potential interventions may be unknown. In this case, the fraction of exposures that could come from a specific source (as estimated by various typing procedures) may serve as a useful constraint on the fraction of exposures that could be prevented by controlling that source.

Example: Using Genetic Data to Attribute Campylobacteriosis Cases to Chicken Consumption

Genetic typing data is sometimes used to help estimate the fraction of bacterial illnesses that originate in different food sources. For example, in estimating the fraction of campylobacteriosis cases caused by chicken consumption, one might consider the following data points:

- Nadeau *et al.* (2002) found that "approximately 20% of human *Campylobacter* isolates were genetically related to genotypes found in poultry "
- Hein *et al.* (2003) noted that "A small number of human isolates [11 out of 101] shared PFGE/AFLP types with poultry isolates [sampled at slaughter in

Austria], however, further studies should also focus on the identification of other sources of *C. jejuni* infection in humans."

- Moore *et al.* (2003) stated that "Human campylobacteriosis is currently the most common cause of acute bacterial gastroenteritis on the island of Ireland… It was the aim of this study to examine the phenotypic and genotypic relatedness of campylobacters isolated from chickens and humans locally. Sixty isolates were subtyped using phenotyping techniques (biotyping, phage-typing), as well as genotyping techniques (multilocus enzyme electrophoresis (MEE), ribotyping) and the data compared. The frequency of shared phenotypes and genotypes between poultry and humans varied depending on the typing technique employed ranging from 98.2% of human isolates sharing a similar resistotyping (MAST) disc type with poultry strains to 20% similarity with MEE typing."

Interpreting such genotype data in terms of sources of exposure is problematic. For example, not all genotypes shared between species necessarily result from one eating the other. (Thus, lambs and chickens, and also people and dogs, have overlapping *Campylobacter* genotypes, but in each pair, neither eats the other.) Schouls *et al.* (2003) caution that "We conclude that typing of *Campylobacter* strains is useful for identification of outbreaks but is probably not useful for source tracing and global epidemiology because of carriage of strains of multiple types and an extremely high diversity of strains in animals." However, the 20% number from Nadeau *et al.* and Moore *et al.* might be viewed as a conservative estimate of the fraction of human isolates that are likely to have come from eating chicken. (An unknown part of this 20% may be due to sources such as contaminated water that are common to chickens, humans, dogs, lambs, and other species.)

Example: Exposure and Hazard Information from Resistance Data

Problem: Suppose that 80% of isolates of a particular bacterium, B, taken from chickens are resistant to antibiotic A; while 10% of isolates from human patients infected by bacterium B are resistant to antibiotic A. How, if at all, do these observations constrain the fraction of human patients infected by bacterium B that might be caused by chicken? Are they consistent with a belief that most human cases of A-resistant B infections come from eating chicken?

Solution: On the face of it, if it were true that "most" (e.g., 50% or more) of human cases of A-resistant B infections came from eating chicken, then at least (50% infections from chicken) × (80% A-resistance rate in chicken-borne B) = 40% of human infections should be resistant. Since only 10% actually are resistant, this is evidence against the belief that most A-resistant cases of B come from eating chickens (assuming that the 80% and 10% numbers are reliable.) However, this

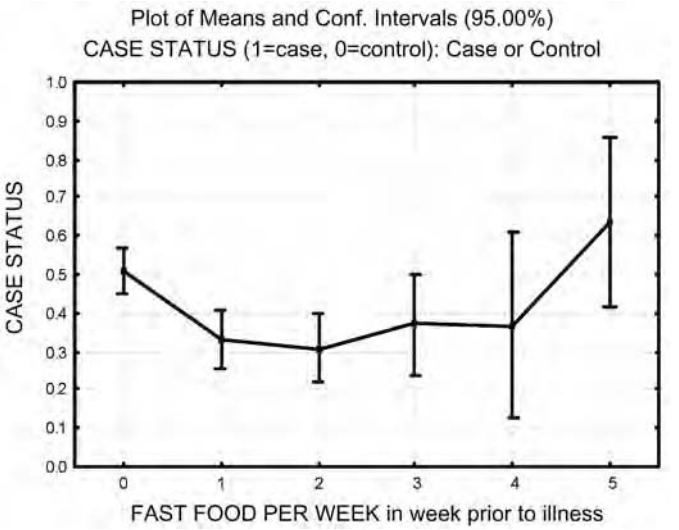
reasoning makes an implicit additional hypothesis: that all bacteria of type B can be treated as if they were homogeneous or interchangeable with respect to causation of human illnesses. If this hypothesis is abandoned, then the stated facts pose no useful constraints on the fraction of human cases of A-resistant B that come from chickens. For example, suppose that there is a particularly virulent strain of bacterium B, call it B*, that accounts for *all* cases of A-resistant B infections in humans, and that comes only from chickens. Then the belief that most A-resistant cases of B come from eating chickens would be correct after all (all of them do!) and would be entirely consistent with the data presented. Thus, while typing information can potentially inform hazard identification (could chicken be the main source of A-resistant B infections in people?) and exposure assessment (no more than 12.5% of human exposures can come from chickens if human isolates have a 10% resistance rate and chickens have an 80% resistance rate, since 80% × 12.5% = 10%), such arithmetical inferences require microbiological justification, such as knowledge that all strains are approximately equally virulent or that resistance rates in isolates from human patients accurately reflect resistance rates in the causes of infection and are not caused by something else, such as taking antibiotics in conjunction with chicken (perhaps in the form of chicken soup) in the early stages of illness.

Example: Using Epidemiological Data to Attribute Campylobacteriosis Cases to Chicken Consumption

Epidemiological data can complement genotyping data and other biological typing data, such as serotypng, resistance typing, and phenotypic typing information, in estimating the fraction of bacterial illness cases resulting from a specific source of exposure. Continuing the example of estimating the fraction of campylobacteriosis cases caused by eating chicken:

- A prospective case-control study from Quebec (Michaud *et al.*, 2001) identified poultry as the "principal suspected source of infection" in about 10% of cases, comparable to drinking tap water at home (9%). However, not all cases had identified probable sources causes in this study.
- Our analysis of data of Kapperud *et al.* (2003) from Norway suggests that a fraction 7/211 = 3.3% of all cases might be due to eating undercooked poultry, after adjusting for other variables by conditioning using classification tree analysis. (Eating undercooked poultry is associated with eating other undercooked meats, so some of the excess risk associated with eating undercooked poultry may be caused by other factors.)
- Using a logistic regression model, Friedman *et al.* (2004) estimated higher population attributable fractions (PAF = 24%) for eating chicken prepared in restaurants. However, this model did not distinguish between association and causation and no corrections for model uncertainty were reported, nor were high-order interactions among variables included. The following plot of some of the data analyzed in that study shows the fraction of subjects who had eaten

in a fast food restaurant 0, 1, 2, 3, 4, or 5 times in the past week who ended up as cases (campylobacteriosis cases) rather than as controls.



Its U shape suggests that the relation between chicken consumption (meals per week) and risk of campylobacteriosis is nonmonotonic, in contrast to the logistic regression model's assumption of monotonicity. Thus, the validity of the logistic regression model, and of PAF calculations based on it, is unknown.

If plausible estimates of the fraction of campylobacteriosis cases caused by eating chicken fall roughly in the range from 3% to 24%, then a point estimate of about 9% with an uncertainty factor of about 3 (i.e., from 0.03 to 0.027) may be realistic.

A quick sanity check that does not use statistical models is based on the observation that *Campylobacter* levels in processed broiler carcasses fell by about 90% since the mid-nineties (Stern and Robach, 2003). Assuming a proportional reduction in human risk of chicken-borne campylobacteriosis, the true fraction of campylobacteriosis cases caused by eating chicken could have fallen from a pre-1995 value of at most 100% to a current value of at most 10%. This is an admittedly crude calculation, based on a single study, but it suggests that an estimate of 10% or less of human campylobacteriosis cases currently caused by eating chicken may not be unreasonable. Table 2 list some other possible sources of campylobacteriosis, as discussed further in the next example.

Example: Biology and Epidemiology of Protective Exposures

If campylobacteriosis cases do not primarily some from eating chicken, as the previous results and others (e.g., Michaud *et al.*, 2004) suggest, then where do they

Table 2: Risk Factors and Sources Associated with Campylobacteriosis

Source/risk factor	Study
Bird-pecked milk bottle tops	Lighton *et al.*, 1991. Note: "Previously recognised associated factors such as outdoor activities, pet ownership and consumption of chicken showed no significant association."
Children	Skirrow, 1977. "Spread of infection was observed within 12 out of 29 households, and in these cases children were usually implicated."
Chicken	Norkrans and Svedhem 1982; Rosenfield *et al.*, 1985;Eberhart-Phillips *et al.*, 1999; Studahl and Andersson, 2000; many others in many countries. See also restaurant chicken.
Contact with dogs, puppies with diarrhea	Skirrow, 1977; Blaser *et al.*, 1980; Brieseman, 1990
Cow	Stalder *et al.*, 1983. Outbreak traced to cow. Finch and Blake, 1985.
Farm visit in previous 2 weeks	Gillespie *et al.*, 2003
Food handler	Olsen *et al.*, 2000 (shedding, contaminated pineapple and gravy)
Food handling	Brown *et al.*, 1988 "A point-source outbreak or *Campylobacter* infection affected 24 of 51 delegates attending a business lunch. …Cross-contamination as a result of handling raw and cooked food consecutively was a possible cause of the outbreak"
Foreign travel	Rodrigues *et al.*, 2001, many other studies in many countries
Hamburger (raw)	Oosterom *et al.*, 1980. Probable cause of barrack outbreak
Milk (unpasteurized)	Gillespie *et al.*, 2003, Korlath *et al.*, 1985; Finch and Blake, 1985; many other studies since early 80s
Organic meats at home in winter	Gillespie *et al.*, 2003
Pets with diarrhea	Gillespie *et al.*, 2003
Pork (undercooked)	Kapperud *et al.*, 2003 for Norway
Restaurant food, commercial catering establishments	Gillespie *et al.*, 2003; Frost *et al.*, 2002 for outbreaks in England and Wales, 1995-9
Restaurant chicken	Rodrigues *et al.*, 2001; Eberhart-Phillips *et al.*, 1999; many others in multiple countries.
Salad	Blaser *et al.*, 1988
Salad bar/food preparation and storage in the facility kitchen	Kirk, 1997. "In seven affected groups of people using the facility, the attack rate ranged between 19% and 67%."
Tuna salad at a summer camp	Roels *et al.*, 1998
Sex between men	Gaudreau and Michaud, 2003 for resistant cluster in Montreal
Water from private water supplies	Said *et al.*, 2003 for outbreaks in England and Wales, 1970-2000; Pebody *et al.*, 1997 attribute 6 out of 21 outbreaks in England and Wales from 1992-4 to private water supplies
Water (undisinfected)	Kapperud *et al.*, 2003. "Drinking undisinfected water, reported by 53% of cases, was a leading risk factor in this study. Drinking water may constitute the common reservoir linking infection in humans and animals, including poultry and wild birds."
Cabbage stew with beef	Steffen *et al.*, 1986. Outbreak in German school children

come from? Although the epidemiology of campylobacteriosis outbreaks is only partly relevant for understanding the epidemiology of sporadic cases, it can help to identify potential sources of exposure, especially because the origins and/or vehicles by which pathogens are transmitted to outbreak victims are often traced with much greater certainty than is possible for any individual case.

Table 2 lists some sources and risk factors for campylobacteriosis identified in previous literature from several countries in studies of both outbreaks and sporadic cases. The sources may be roughly grouped as: restaurant-related (including cafeterias, commercial catering and food establishments), farm and animal related, foreign travel-related, water-related, and other. Chicken at restaurants is implicated as a source of sporadic cases (Rodrigues *et al.*, 2001) as well as of outbreaks, as is other restaurant food (Frost *et al.*, 2002). Chicken prepared and eaten at home is *not* associated with any increased risk in several studies (Lighton *et al.*, 1991, Rodrigues *et al.*, 2001), but is associated with a large (e.g., 40%), statistically significant reduction in risk in several studies (e.g., data of Friedman *et al.*, 2000; Effler *et al.*, 2001). This contradicts the widespread and oft-repeated misconception that "Most [*Campylobacter*] infections are acquired by the consumption and handling of poultry" (Allos, 2001), or that "In industrialised countries, most infections are acquired through the handling and consumption of poultry meat." (Butzler, 2004). A possible explanation is that eating poultry and other meats in some commercial food establishments with unhygienic kitchen practices may increase risk of campylobacteriosis; while eating them at home can build protective immunity and help to protect against campylobacteriosis.

The hypothesis that low (sub-infectious) levels of exposure to *Campylobacter* in chicken builds immunity can potentially be supported or refuted by testing it for other foods. The importance of individual immunity in mediating susceptibility to campylobacteriosis has previously been noted for outbreaks following consumption of raw milk. For example, in one study:

> "22 (88%) of 25 students who consumed raw milk for the first time became infected…. Among ten persons who chronically consumed raw milk, none was ill, a striking difference from the 76% attack rate among the 25 acutely exposed students. The quantity of raw milk consumed was directly related to the occurrence and severity of illness. Acutely infected students showed significant rises in *C jejuni*-specific immunoglobulins, whereas the low antibody levels seen in unexposed persons did not rise. In contrast, acute-phase serum samples from persons with chronic exposure to raw milk showed elevated antibody levels to *C jejuni*. These findings indicate that chronic raw milk consumption is associated with elevated levels of *C jejuni*-specific serum antibodies and with immunity to symptomatic infection." (Blaser *et al.*, 1987; see also Walz, 2001.)

Conversely, a compromised immune system is a significant risk factor for campylobacteriosis: "The average annual incidence of *Campylobacter* among AIDS cases (519/100,000) exceeded the crude population rate by 39-fold and exceeded the rate among males aged 15-55 years by 35-fold" (Sorvillo *et al.*, 1991). Demographic attributes significantly associated with increased risk of campylobacteriosis include being male, being an infant or in one's twenties, and ethnicity (*Campylobacter* sentinel surveillance scheme collaborators, 2003).

In deciding what fraction of campylobacteriosis cases should be attributed to any specific cause such as consumption of undercooked chicken in restaurants, it is necessary to decide how to account for: (a) *Protective effects* (e.g., of foods that build acquired immunity); and (b) *Interactions* between host and exposure in determining risk. For example, should the risk of campylobacteriosis for AIDS patients who eat undercooked meat be attributed entirely to the meat, or somewhat to the undercooking and/or to compromised immunity? Such policy decisions have seldom been addressed explicitly in risk attribution discussions.

One approach to solving (or bypassing) these difficulties, taken in this book, is to focus on the human health consequences caused by specific interventions. From this standpoint, the fraction of all campylobacteriosis cases caused by eating chicken, for example, is of limited interest. What matters for practical decision-making is the *change* in campylobacteriosis rates and their health consequences (e.g., illness-days per year) that would result if specific interventions (such as banning enrofloxacin) were made. The fraction of cases per year that would be prevented by a specific intervention can be estimated by predictive risk models (e.g., based on simulation or systems dynamics models), without resolving all of the more philosophical quandaries of attribution. Hence, we focus on the question of what fraction of annual cases could be prevented by specific interventions, according to specified models of the causal relation (possibly U-shaped) between exposure and risk, and do not address other questions about attribution of risk. (However, for an axiomatic framework for attribution of risk in the presence of joint, interacting causes, based on the Shapley value for cost allocation games, see Gefeller *et al.*, 1998 and Cox, 2001, Chapter 4. A key insight from this more formal approach is that the proportion of *risk* that should be attributed to a source is often just the proportion of *exposure* that comes from that source. This is true even when there are strong nonlinearities in the dose-response function, e.g., a response threshold, provided that exposures from different sources affect risk symmetrically.)

Example: Calculating a Preventable Fraction with a Protective Effect

Problem: The risk of a person becoming ill from ingesting 10^x CFUs of a bacterium B in a certain time interval (e.g., in one meal or within a certain number of hours or days, depending on the biology of infection) is as follows: $r(x) = 0.1x$ for $1 \le x \le 4$; $r(x) = 0.4$ for $x > 4$; $r(x) = 0.2 - 0.1x$ for $0 \le x \le 1$. Before a risk management intervention is implemented, the amount of exposure received in each time interval is uniformly distributed between 0 and 4: $x \sim U[0, 4]$. What fraction of illnesses are prevented by an intervention that reduces all exposures by 50%?

Solution: In general, if the cumulative distribution function for exposure is $\Pr(\text{exposure} \le x) = F(x)$ and the conditional probability of illness when exposure is x is r(x), then the risk induced by exposure distribution F and dose-response function r(x) is, in applied probability notation: risk $= E_F[r(x)] = \int r(x)dF(x)$. In this example, prior to the intervention, the expected number of illnesses per unit time is:

risk before intervention
$= \Pr(0 \le x \le 1) \times E[r(x) \mid 0 \le x \le 1] + \Pr(1 \le x \le 4) \times E[r(x) \mid 1 \le x \le 4]$
$= 0.25 \times 0.15 + 0.75 \times 0.25 = 0.225.$

Following the intervention, with all exposures reduced by half, the risk is:

risk after intervention
$= \Pr(0 \le x \le 0.5) \times E[r(x) \mid 0 \le x \le 0.5] + \Pr(0.5 \le x \le 1) \times E[r(x) \mid 0.5 \le x \le 1]$
$+ \Pr(1 \le x \le 2) \times E[r(x) \mid 1 \le x \le 2]$
$= 0.25 \times 0.175 + 0.25 \times 0.125 + 0.5 \times 0.15 = 0.15.$

Thus, the preventable fraction for this intervention is $[(0.225 - 0.15)/0.225] = 1/3$. That the preventable fraction is less than the fractional reduction in exposure accomplished by the intervention (namely, 50%) is explained by the nonlinear (V-shaped) dose-response function.

4.1 Attribution through Multiple Stages

Attribution of measured health outcomes to the exposure sources that contribute to them may be repeated, driving the attribution back recursively to successively earlier stages in the chain or network of bacterial sources that ultimately lead up to the exposures that caused harm. Such calculations work backward from clinical outcomes to intermediate sources (e.g., community *vs.* nosocomially acquired sources), and then to their predecessors (e.g., food or water consumption, contact with infected or contaminated animals and humans, etc.), and so forth. The process is repeated, ultimately leading back to a fraction of microbial load (resistant,

susceptible, or both) that is estimated to be due to a particular source of interest, such as antibiotic-resistant bacteria in animals, selected by use of antimicrobial agents on the farm. At each stage, the fraction of load contributed by each preceding source is estimated. The fractions along the path leading from adverse outcomes back to the ultimate source of interest, such as antimicrobial use on the farm, are then multiplied and the results are interpreted as estimates of the fraction of cases per year that could be prevented by eliminating any of the steps along the path, including antimicrobial use on the farm.

When multiple paths lead from hazards affected by a risk management intervention (such as bacteria in food or animal antibiotic use on the farm) to exposures that cause measured adverse health consequences (such as cases of illnesses or resistance in human patients), then the contributions to human health impacts from different paths are summed to obtain the total impact of the intervention. Thus, just as for simulation-based approaches that work "forward" from sources to effects, attribution-based models that work "backward" from effects to their causes may lead to the same type of algebraic formula, expressing risk as a sum of products of factors. However, the logic of the approach is different. By starting with cases, the attribution approach estimates the fractions of exposures that actually result in (or at least co-occur with) illnesses (i.e., cases) that come from different sources.

Example: Two-Stage Preventable Exposure Fractions

Suppose that 10% of cases of bacterial illness B in humans are attributed to chickens (meaning that removing chickens as a source would prevent 10% of cases) and that 20% of bacteria B in chickens are attributed to contamination spread by flies. Then 10% × 20% = 2% of human cases would be attributed to contamination spread to chickens by flies. If flies also spread bacterium B from chickens to other reservoirs and hosts (e.g., family pets), some of which then lead to illness cases in humans, then a Leontief input-output matrix style of analysis can be used to define the total fraction of human cases to attribute to each source, taking both direct and indirect contributions along all paths into account. This is illustrated in the next example.

Example: Matrix Calculation of Preventable Exposure Fractions

Setting: Consider an ecosystem in which pathogenic bacteria move among three compartments: humans (H), water (W), and chickens (C). The direct effects of the different sources in contributing bacteria to each other are described by the following system of equations:

$$H = 0.5W + 0.1C + 0.4$$
$$W = 0.2H + 0.2C + 0.6$$
$$C = 0.3W + 0.7$$

The coefficients (including the intercept) in each term sum to 1 and the initial values of H, W, and C are normalized to 1 (i.e., each is at 100% of its baseline equilibrium value). The intended interpretation of these equilibrium equations is that 50% of human exposure to the bacterium comes via water, 10% from chicken, and the remaining 40% from outside the system. Similarly, 20% of the bacteria in water come from humans (e.g., via sewage), 20% from chickens, and 60% from outside the system. 30% of the bacteria in chickens come from water and the rest from outside the system.

Problem: What fraction of human exposure to the bacteria would be prevented if a new technique of animal husbandry eradicated the contribution from chicken?

Solution: Setting C = 0 makes the third equation irrelevant and reduces the system to the following:

$$H = 0.5W + 0.4$$
$$W = 0.2H + 0.6$$

This is of the form $x = Ax + b$, with solution $x = (I - A)^{-1}b$, where I is the 2 x 2 identity matrix, b = [0.4, 0.6]' is the column vector of exogenous levels, and A is the matrix of coefficients A = [0, 0.5 ; 0.2, 0] (using the MATLAB™ convention that rows are separated by semicolons.) The solution can be found by MATLAB™ by entering the following model equations, which can be adapted to large models:

```
A = [0, 0.5 ; 0.2, 0];
I = [1, 0; 0, 1];
b = [0.4, 0.6]';
x = (inv(I – A))*b
```

The result returned in this case is: x = [0.7778, 0.7556]', i.e., the new equilibrium level of exposure in the human compartment is 0.7778 = 7/9. Removing chicken as a source prevents 22.22% = 2/9 of human exposures.

The technique illustrated in this example is appropriate for calculating changes in equilibrium exposures in systems described by linear equations. If the removal or reduction of one source of bacteria leads to growth of another, however, then simple linear analysis of equilibrium fractions no longer suffices, and predictive modeling must be used to calculate changes in human exposures.

In attribution exposure models, uncertainties about the attributable fractions at different stages are commonly treated by using upper-bound estimates. If the product of the upper bound estimates is small, then the true but unknown value of the product is also small, and this information may be sufficient to support a decision that no intervention is required. If the product of the upper bound estimates is large enough so that this conclusion cannot be justified, then the uncertainty analysis can be refined by estimating probability distributions for the fractions at different stages and applying basic Monte Carlo uncertainty analysis (see e.g., Vose, 2000) to obtain the probability distribution for their product.

5. SYSTEMS DYNAMICS MODELS OF EXPOSURE

For antimicrobial risk assessments, including risk assessments of animal antibiotic uses, the exposures of concern specifically include exposures to antibiotic-resistant bacteria. Exposure assessment requires estimating the changes in human exposures to resistant and susceptible bacteria that would be caused by alternative risk management interventions. This section illustrates how such estimation can be accomplished, using the model introduced in Chapter 3 for purposes of illustration.

5.1 Model Development and Simplifying Assumptions

The equations of the model are as follows:

$$IA = IA_0 + A^*(IA_1 - IA_0) \quad (1)$$

$$dIH/dt = [a_1 + b_1(1 - IA) + c_1IA](1 - IH) - r_1IH \quad (2)$$

$$dRA/dt = [a_2 + b_2A + d_2IH](1 - RA) - r_2RA \quad (3)$$

$$dRH/dt = [a_3 + d_3IH + fRA](1–RH) - r_3RH \quad (4)$$

where IA = the ill animal fraction (e.g., the fraction of chicken servings from airsacculitis- or necrotic enteritis-positive flocks), IH = ill human fraction (e.g., the fraction of the population that has the bacterial illness of concern at any time), RA = resistant fraction of bacteria from animals, and RH = resistant fraction of bacteria from humans. The controllable input is A, with possible values of 1 if animal antibiotic use continues and A = 0 otherwise. Equation (1) shows that the effect on animal health of a ban on animal antibiotic use is a change in the prevalence of ill animals at slaughter from IA_1 to $IA_0 > IA_1$.

The coefficients in this model can be estimated from empirical information. To illustrate principles as clearly as possible with minimum algebraic manipulation, we will focus on comparing the steady-state

equilibrium exposure levels before and after an intervention replacing A = 1 (continued use) with A = 0 (discontinued use). The following assumptions are made for simplicity. (Subscripts of 1 and 0 denote the equilibrium values of dynamic variables when A = 1 and when A = 0, respectively, i.e., with and without continued animal antibiotic use.)

- The pre-intervention levels IA_1, IH_1, RA_1, and RH_1 are all known, e.g., from surveillance studies in animals and humans.
- $IA_1 = 0$, or is close enough to zero to be negligible. This is appropriate for animal illnesses such as airsacculitis or necrotic enteritis that have been effectively controlled.
- *All* animal resistance is caused by animal antibiotic use and would eventually disappear if such use were discontinued (i.e., if A were set to zero. Thus, $RA_0 = 0$.) Mathematically, this implies that $a_2 = d_2 = 0$ in equation (3). This extreme assumption provides a bounding case that can be relaxed in sensitivity analyses or in more detailed analyses.
- Similarly, *all* human antibiotic resistance originates in resistant bacteria from food animals, so that if RA = 0, then eventually RH = 0.

With these assumptions, the model equations reduce to the following:

$$IA = (1 - A)*IA_0, \text{ for } A = 0 \text{ or } 1. \quad (1a)$$

$$dIH/dt = [a_1 + b_1*(1 - IA) + c_1*IA]*(1 - IH) - r_1IH \quad (2a)$$

$$dRA/dt = b_2*A*(1 - RA) - r_2*RA \quad (3a)$$

$$dRH/dt = f*RA*(1–RH) - r_3*RH \quad (4a)$$

In steady-state equilibrium, all time derivatives become zero and equation (2) simplifies to: $[a_1 + b_1(1 - IA) + c_1IA](1 - IH) = r_1IH$. Solving this for the equilibrium value of IH_0 yields:

$$IH_0 = [a_1 + b_1*(1 - IA_0) + c_1*IA_0]/[r_1 + a_1 + b_1*(1 - IA_0) + c_1*IA_0]$$
$$= 1/[1 + r_1/(a_1 + b_1*(1 - IA_0) + c_1*IA_0)] \quad (2b)$$

Substituting $IA_1 = 0$ into the equilibrium equation simplifies it to:

$$IH_1 = [(a_1/b_1) + 1]/[(a_1/b_1) + 1 + (r_1/b_1)] = 1/[1 + r_1/(a_1 + b_1)] \quad (2c).$$

Equations (3a) and (4a) are irrelevant for steady-state equilibrium analysis, as we assume that RA_1 and RH_1 are known and that $RA_0 = RH_0 = 0$.

5.2 Estimating Model Parameters from Data

To make predictions from a parametric model such as (2b), it is usual to first estimate the model parameters from what is known and observed and then to use them to predict model outputs for new inputs. The model parameters can be estimated from the following empirical inputs:

1. *How many times more likely to cause illness is a serving from an ill animal than a serving from a well animal?* This determines the ratio c_1/b_1. Limited empirical data suggest that processed carcasses from airsacculitis-positive chicken flocks may carry an average of about 10 times more *Campylobacter* than processed carcasses from airsacculitis-negative flocks (due largely to higher processing error rates and fecal contamination among underweight birds), although there is considerable flock-to-flock heterogeneity (Russell, 2003). If risk of illness is proportional to microbial load on processed carcasses (presumably attenuated for each serving by a product of random reduction factors for subsequent handling, processing, storage, and preparation), then an estimated value for the c_1/b_1 ratio of about $c_1/b_1 = 10$ might be plausible as a base case, i.e., as a starting point for uncertainty and sensitivity analyses. (Non-proportional dose-response relations are considered in Chapter 8 and its appendix on the log-exponential model.)
2. *What proportion of human illnesses are caused by the food animal or commodities of concern*, as opposed to other sources (e.g., vegetables, drinking water, pets, etc.?) In general, this proportion determines the ratio $[b_1(1 - IA) + c_1IA]/[a_1 + b_1(1 - IA) + c_1IA]$. For $IA_1 = 0$, however, this ratio simplifies to: $b_1/(a_1 + b_1)$. Hence (on dividing numerator and denominator by b_1), this proportion determines the ratio a_1/b_1. For example, if 10% of human campylobacteriosis illnesses are caused by consumption of chicken in the base case, then $b_1/(a_1 + b_1) = 10\%$ and so $b_1 = 0.1 \times (a_1 + b_1)$, $0.9b_1 = 0.1a_1$, and $a_1/b_1 = 9$. If instead 60% of human campylobacteriosis illnesses are caused by consumption of chicken (the base case assumed by FDA-CVM, 2001 was about 57%), then $b_1/(a_1 + b_1) = 0.6$, so $b_1 = 0.6 \times (a_1 + b_1)$, $0.4b_1 = 0.6a_1$, and $a_1/b_1 = 0.67$. We will use both values, $(a_1/b_1) = 9$ and $(a_1/b_1) = 0.67$, in subsequent calculations to illustrate the sensitivity of results to this ratio.
3. *What is the mean duration of illness?* The reciprocal of this number determines the recovery rate, r_1. For example, if the mean duration of illness is about 6 days, then $r_1 = (1/6)$ per day = 5 per month. If the duration of illness is significantly different in some subpopulations, or for resistant compared to susceptible illnesses, then the model could be fit separately to each such subpopulation, but for purposes of illustration,

a single recovery rate is assumed here, implying an exponentially distributed random recovery time for individuals.

4. *What is the initial endemic illness rate in humans* prior to intervention? For example, for campylobacteriosis, there are about (13.4E-5 reported cases per capita-year) × (38 assumed cases per reported case, from Mead *et al.*, 1999) = 0.0051 cases per capita-year. Assuming an average of 6 days per illness, each illness causes an average of (6/365) of a year to be spent ill, giving an illness fraction (i.e., the probability that a randomly selected member of the population on a randomly selected day is ill) of: IH_1 = (0.0051 cases per capita-year) × (6 days per illness/365 days per year) = 8.4E-5, prior to any intervention.
5. *If the animal antibiotic use is discontinued (A is set to zero), by how much will the ill animal fraction increase?* This requires assessing how effective the animal antibiotic is in reducing animal illnesses (or in reducing the prevalence of underweight animals going to slaughter, if that is the only effect of antibiotic use on animal health. Recall that IA describes the fraction of animals contributing to servings with significantly increased microbial loads, whether the detailed cause is a specific bacterial illness or simply being underweight and at increased risk of processing errors and fecal contamination.) The answer to this question determines ($IA_0 - IA_1$) in general, and IA_0 in the particular case $IA_1 = 0$. For purposes of illustration, we will consider an increase in the animal illness fraction from approximately 0% to approximately 1%. In reality, the endemic rates of animal illnesses such as necrotic enteritis or airsacculitis increase by different amounts in different countries (see e.g., VLA, 2004 for necrotic enteritis rates in the U.K. following a ban on virginiamycin and other growth promoters). It is therefore useful to calculate the exposure consequences of a withdrawal on the basis of an increase of only 1%, and then multiply the results by the estimated true percentage increase in animal illnesses if A is set to 0.

These empirical inputs allow the parameters in equation (2) to be fully specified. When A = 1 (continued animal antibiotic use) and $IA_1 = 0$, equation (2) determines the equilibrium endemic value of IH_1 via the equation: $IH_1/(1 - IH_1) = (a_1 + b_1)/r_1$, which for a small IH_1 (such as IH_1 = 8.4E-5) gives answers that are numerically almost identical to $IH_1 = (a_1 + b_1)/r_1$. Substituting the estimates $r_1 = 5$ per month and $a_1 = 9b_1$ from above into this formula yields: $IH_1 = 10b_1/5 = 2b_1$, or $b_1 = IH_1/2$ = 4.2E-5 per month. Then $a_1 = 9b_1 = 9 \times 4.2E\text{-}5 = 3.78E\text{-}4$ and $c_1 = 10b_1 = 10 \times 4.2E\text{-}5$ = 4.2E-4, using the preceding suggested base case estimates of these parameter ratios. So, the fully specified version of equation (2) using base case parameter estimates (and a time scale measured in months) becomes:

$$dIH/dt = b_1[(a_1/b_1) + (1 - IA) + (c_1/b_1)IA](1 - IH) - r_1IH$$
$$= (4.2E\text{-}5)*[9 + (1 - IA) + 10*IA]*(1 - IH) - 5*IH.$$

This is a linear ODE with constant coefficients. It can readily be solved to answer questions about how changing IA will affect IH.

If 60% (rather than 10%) of human illnesses stem from animal bacteria, so that $(a_1/b_1) = 0.67$, then the equation $IH_1 \approx (a_1 + b_1)/r_1$ implies that $1.67b_1 = r_1*IH_1$, or $b_1 = [(r_1*IH_1)/1.67] = (5*8.4E\text{-}5)/1.67 = 2.515E\text{-}4$. The model parameter estimates are then: $[a_1, b_1, c_1, r_1] = [0.67 \times 2.515E\text{-}4, 2.515E\text{-}4, 10 \text{ x } 2.515E\text{-}4, 5] = [1.685E\text{-}4, 2.515E\text{-}4, 2.515E\text{-}3, 5]$.

5.3 Using the Model to Make Predictions

For comparing the pre-intervention and post-intervention equilibrium values, equation (2b) can be used to quantify IH_0, which is then compared to the baseline value of $IH_1 = 8.4E\text{-}5$. If illnesses are proportional to the exposures to the bacteria that cause them, then the fractional change in illnesses, $(IH_0 - IH_1)/IH_1$, will be the same as the fractional change in exposures. [If resistant and susceptible bacteria cause different durations of illness, then separate values of r_1, say r_{1s} and r_{1r}, could be assessed and illnesses due to each type of bacterium could be tracked separately. Here, we continue to assume that susceptible and resistant strains have roughly the same average recovery rate, as appears plausible for domestically acquired campylobacteriosis cases and antibiotics such as macrolides and fluoroquinolones (see e.g., Ang and Nacham, 2003).]

Example: Calculating Exposure Consequences of a Ban

Problem: In the preceding model, suppose that the animal antibiotic has been used (A = 1) for long enough so that the system has reached steady-state equilibrium. If the animal antibiotic use is now discontinued (setting A = 0), what will be the resulting relative change in the new equilibrium exposure and illness rates? Answer for 10% and for 60% of human illnesses IH caused by food animal bacteria.

Solution: By hypothesis, changing A from 1 to 0 increases the ill animal fraction from $IA_1 = 0$ to $IA_0 = 0.01$. The new equilibrium IH_0 is given by equation (2b):

$$IH_0 = [a_1 + b_1*(1 - IA_0) + c_1*IA_0]/[r_1 + a_1 + b_1*(1 - IA_0) + c_1*IA_0] \quad (2b)$$

Substituting the rough estimates $r_1 = 5$, $b_1 = 4.2E\text{-}5$, $a_1 = 9*b_1$ and $c_1 = 10*b_1$ yields $IH_0 = 8.476E\text{-}5$. Entering the following lines in MATLAB™:

r1 = 5; b1 = 4.2E-5; a1 = 9*b1; c1 = 10*b1; IA0 = 0.01; IH1 = 8.4E-5;
IH0 = [a1 + b1*(1 - IA0) + c1*IA0]/[r1 + a1 + b1*(1 - IA0) + c1*IA0]
FractionalChange = $(IH_0 - IH_1)/IH_1$

returns the following results:

IH0 = 8.4749e-005 FractionalChange = 0.0089.

Thus, the human illness rate (and the corresponding exposures) are predicted to increase by only 0.89% for a 1% increase in animal illness rates.

If the fraction of human bacterial illnesses from food animals is assumed to be 60% rather than 10%, then the preceding calculations are revised as follows:

r1 = 5; b1 = 2.515E-4; a1 = 1.685E-4; c1 = 2.515E-3; IA0 = 0.01; IH1 = 8.4E-5;
IH0 = [a1 + b1*(1 - IA0) + c1*IA0]/[r1 + a1 + b1*(1 - IA0) + c1*IA0]
FractionalChange = $(IH_0 - IH_1)/IH_1$

resulting in:

IH0 = 8.8519e-005 FractionalChange = 0.0538.

In this case, a 1% increase in the ill animal fraction induces a predicted increase of about 5.38% in human illnesses and in the corresponding bacterial exposures that cause them. In this scenario, a 20% increase in the ill animal fraction (see VLA, 2004 for necrotic enteritis increase data) would more than double the baseline human exposure levels and illness rates.

The systems dynamics approach illustrated in this section can be extended to more complex causal models in which not all human resistant isolates necessarily come from resistant bacteria in food animals and not all resistant isolates in animals necessarily come from animal antibiotic use. The basic steps of equation development, parameter estimation from data, and model-based prediction remain the same, but additional empirical questions must be answered to obtain needed inputs. (For example: How quickly, e.g., with what half-life, will resistance to the antibiotic decrease in animals if use is terminated? The answer determines the value of r_2.)

Systems dynamics models can also be used to calculate (or simulate) time courses of dynamic variables following a change in controllable inputs. For example, solving equation (2) numerically or symbolically shows how quickly IH(t) increases following termination of animal drug use. Solving the system of equations (1)-(4) shows how each of its variables changes over

time following the change in animal drug use. However, for many risk management policy analyses, comparison of the steady-state equilibria before and after intervention provides the most essential information to support decision-making. Chapter 8 develops such comparisons further.

6. EXPOSURE ASSESSMENT METHODS TO AVOID

The following approaches to exposure assessment seek to simplify the process of exposure assessment by ignoring microbial load information. In general, they give inaccurate results and should not be used.

6.1 Prevalence-Based Exposure Metrics

Models that only use dichotomous exposure summaries (e.g., "contaminated" vs. "not contaminated" for chicken carcasses or servings) do not contain sufficient information to allow confident, accurate risk predictions when risk depends on *how much* exposure occurs, and not simply on *whether* it occurs. (For example, multiplying all exposures by 1,000,000 might well increase risk, but it would not affect prevalence.) Such summaries of exposure prevalence should not be used for risk assessment unless quantitative information does not significantly affect risk. However, using prevalence together with a conditional probability distribution for microbial load given that it is greater than zero is perfectly acceptable. The WHO/FAO 2002 risk assessment for *Salmonella* illustrates this approach in detail (*op cit,* http://www.who.int/foodsafety/publications/micro/Salmonella/en/).

Technical Note: Discretizing Continuous Exposures. If a qualitative or categorical summary of exposure is desired, then histogram-fitting methods are available that lose less relevant information than prevalence measures and that can allow useful predictions. For example, classification tree algorithms (Zhang and Singer, 1999) automatically bin more detailed exposure-response data into a few contiguous aggregate exposure intervals (or combinations of intervals, if there are multiple exposure factors) that predict similar response probability levels.

Example: Prevalence Does Not Predict Risk

The WHO/FAO 2002 risk assessment for *Salmonella* reported that:

> "The effect was assessed of reducing the numbers of *Salmonella* on poultry carcasses without changing the prevalence of contaminated carcasses. The values of the cumulative concentration distribution used in the baseline scenario were

> reduced by 50% (approximately 0.3 logCFU per carcass…). The model was run using the reduced level of contamination while maintaining the prevalence at 20% and with no changes in any of the other parameters. Unlike a change in prevalence, a change in concentration of the pathogen does not necessarily have a linear relationship with the risk outcome. … The servings were estimated to be contaminated and potentially undercooked approximately 2% of the time. That statistic remains unchanged if the level of contamination is reduced. The expected risk per serving, which incorporates the prevalence of contaminated servings and the probability of undercooking, was estimated to be 1.13E-5 (1.13 illnesses per 100,000 servings) in the original case, and 4.28E-6 (4.28 per 1,000,000 servings) in the situation when the level of contamination is reduced. The expected risk per serving is therefore reduced by approximately 62%."

Thus, in this example, a reduction in contamination that reduces risk by 62% has no effect on prevalence of contamination, indicating that prevalence of contamination alone does not determine risk. This point is generally well recognized (see e.g., Ross and McMeekin, 2003; and Rosenquist *et al.*, 2003 for a discussion of similar points for campylobacteriosis), although some previous risk assessments (FDA-CVM, 2001) mistakenly attempted to predict risk from prevalence without using the essential information about the conditional distribution of microbial load.

6.2 Holistic Statistical Exposure Modeling

A common mistake is to create a regression model describing the *statistical* relation between population responses and population exposures (and perhaps other predictors) and then to treat the statistical relation as if it described a *causal* relation that could be used to predict effects of changing the exposure variable. Such aggregate-level statistical modeling does not account adequately for heterogeneity in individual exposure-response relations or distinguish between statistical associations and causal relations that hold at the individual level. It can yield false conclusions about both statistical associations and causal relations among variables (e.g., exposure and response variables) for the individuals in the population. The same caveats hold for aggregate-level statistical models for the relations among exposure variables. For example, consider the following simple linear regression model for estimating exposure levels in food servings (e.g., of chicken) ingested in a population from measured levels of contamination in animals leaving the farm:

$$\text{exposure in ingested servings} = k \times (\text{contamination in animals at farm}).$$

The aggregate parameter k is interpreted as an overall average effective transmission coefficient that tries to account implicitly for the unmodeled

factors intervening between farm contamination and exposures. Such a statistical model has the following limitations that make it unsuitable for causal modeling and health risk assessment (see e.g., Pearl, 2002):

(a) It fails to sum over multiple distinct paths and scenarios, which may be represented by multiple distinct k values for different individuals. The average value of k can be literally meaningless, i.e., it need not give useful information about causal impacts, or even about statistical associations, that hold at the individual level. In general, such "reduced-form" models based on aggregate data may contain biases from omitted variables that create statistical associations very different from causal impacts, and possibly even of opposite sign ("Simpson's paradox", see e.g., Fiedler *et al.*, 2003).

(b) It does not necessarily show how a *change* in the right-side explanatory variables ("contamination in animals at farm", in this example) would affect the left-side variable. (Technically, only *structural equations*, not reduced-form ones, model the causal relations among variables; see Shipley, 2000, Pearl, 2002).

Example: Structural vs. Reduced-Form Equations

To understand why reduced-form models such as simple linear regressions of one aggregate variable against another are inappropriate for risk assessment, consider the following simple example. Suppose that the correct structural equations in a model are:

Exposure ingested = contamination at retail – contamination removed in kitchen (1)

contamination removed in kitchen = (1/3) × contamination at retail (2).

Here, "contamination at retail" is the microbial load on a serving at retail and "contamination removed in the kitchen" is the amount of microbial load that is removed in the kitchen as the serving has been prepared. For simplicity of exposition only, microbial load is modeled here as conserved, i.e., what is ingested is what was on the purchased meat ("contamination at retail") minus what remains in the kitchen ("contamination removed in kitchen"). (In reality, cooking, cleaning, bacterial growth and death, etc. would intervene and microbial load need not be conserved.) In a linear structural equation model such as this one, changes in the variables on the right of each equation are interpreted as affecting the variable on the left, with the coefficients of the right-hand side variables describing the change in the left-hand side variable per unit change in each right-hand side variable (Pearl, 2002). Equation (2) is mathematically (though not causally) equivalent to:

contamination at retail = 3 × contamination removed in kitchen (2').

(This equation does not have an intended causal interpretation because contamination at retail causes contamination in the kitchen, not conversely.) Substituting (2') into (1) gives the *reduced-form* model

Exposure ingested = 2 × contamination removed in kitchen. (3).

This is a simple linear regression model of the form "Exposure = k × contamination" with k = 2. It is valid for some kinds of statistical inference. For example, it correctly predicts that establishments with twice as much contamination in the kitchen will have twice as much exposure ingested under the current conditions described by equations (1) and (2) (since both quantities are proportional to contamination at retail, which may differ across locations.) However, it *cannot* be used to correctly predict the *change* in "Exposure ingested" from an intervention that increases "contamination removed in kitchen" while holding fixed "contamination at retail", e.g., by more thorough cleaning or cooking of meat. It is thus useless (and, indeed, misleading) for predicting the causal impact of such a change, which is what risk managers typically care about. For, equation (1) shows that this causal impact is *negative,* i.e., each unit increase in "contamination removed in kitchen" decreases "Exposure ingested" by one unit, since microbial load left in the kitchen is not on the final serving and hence is not part of the ingested load. But the reduced equation (3) indicates a *positive* statistical relation between them. Equation (1), rather than equation (3), is relevant for predicting causal impacts of interventions.

7. VALIDATING AND REFINING EXPOSURE MODELS

Exposure assessment models should be validated by comparing their predictions under different conditions to measured values of exposures and/or their surrogates. An exposure model can be used to predict microbial loads (or related measured quantities, such as the concentration of CFUs in rinse fluids) at various measurement points and under different conditions, e.g., in different locations, for different seasons, storage and preparation histories, and so forth. Comparing the model-predicted values to measured values using statistical goodness-of-fit tests and diagnostic plots shows whether the observed values are statistically significantly different from the predicted distributions of values. Haas *et al.* (1999), Chapter 6 reviews goodness-of-fit tests for parametric exposure models, (Chapter 7 also discusses validation and uncertainty analysis of simulation models, although with emphasis on dose-response rather than exposure models.)

Example: Validating Predictive Microbiology Exposure Models

The following abstract, from Coleman *et al.* (2003) describes and illustrates efforts to validate (and subsequently improve) predictive models of bacterial growth in food.

> "A novel extension of traditional growth models for exposure assessment of food-borne microbial pathogens was developed to address the complex interactions of competing microbial populations in foods. Scenarios were designed for baseline refrigeration and mild abuse of servings of chicken broiler and ground beef. Our approach employed high-quality data for microbiology of foods at production, refrigerated storage temperatures, and growth kinetics of microbial populations in culture media. Simple parallel models were developed for exponential growth of multiple pathogens and the abundant and ubiquitous nonpathogenic indigenous microbiota. Monte Carlo simulations were run for unconstrained growth and growth with the density-dependent constraint based on the "Jameson effect," inhibition of pathogen growth when the indigenous microbiota reached 10^9 counts per serving. The modes for unconstrained growth of the indigenous microbiota were 10^8, 10^{10}, and 10^{11} counts per serving for chicken broilers, and 10^7, 10^9, and 10^{11} counts per serving for ground beef at respective sites for backroom, meat case, and home refrigeration. Contamination rates and likelihoods of reaching temperatures supporting growth of the pathogens in the baseline refrigeration scenario were rare events. The unconstrained exponential growth models appeared to overestimate *L. monocytogenes* growth maxima for the baseline refrigeration scenario by 1500-7233% (10^6-10^7counts/serving) when the inhibitory effects of the indigenous microbiota are ignored. The extreme tails of the distributions for the constrained models appeared to overestimate growth maxima 110% (10^4-10^5 counts/serving) for *Salmonella spp.* and 108% (6 x 10^3 counts/serving) for *E. coli* O157:H7 relative to the extremes of the unconstrained models. The approach of incorporating parallel models for pathogens and the indigenous microbiota into exposure assessment modeling motivates the design of validation studies to test the modeling assumptions, consistent with the analytical-deliberative process of risk analysis." (*Source:* Coleman *et al.*, 2003)

This example points out both the importance of validating predictive exposure models before using them, by comparing their predictions against observations over a range of values; and also the potential to use such validation studies to improve predictive models by choosing more realistic assumptions when the validation results indicate a significant gap between predicted and observed results.

When predicted exposures do not adequately match validation data, the exposure model should be corrected. This can be done by refining the model to include omitted variables; to more accurately model dependencies among its inputs (Haas, 1999), e.g., by including inhibitory effects of the indigenous microbiota on growth of *L. monocytogenes* in the above example; and/or by using the differences between predicted and observed values to select more appropriate mathematical model forms that can explain and reduce these differences. Advanced statistical methods may help to avoid non-valid models and false predictions. Examples include flexible non-parametric data descriptions and predictions from multiple alternative models that are consistent with available knowledge and data, weighted by their probabilities. (See Cox 2001, Chapter 3 for a survey for risk analysts).

If model-predicted exposures do adequately match validation data according to goodness-of-fit tests and model diagnostics (e.g., plots of residuals), then the exposure model may be used to make predictions for risk assessment within the validated range of conditions. In this case, remaining uncertainty in model parameters, inputs, and predictions should be expressed through confidence intervals for single quantities (e.g., the mean exposure or the upper 95% exposure limit in the population) and through joint confidence regions for multiple quantities, such as the exposures received by different subpopulations. Haas *et al.* (1999) describe and illustrate statistical methods for quantifying uncertainties in microbial risk assessments.

8. QUALITATIVE EXPOSURE ASSESSMENT

Although, as discussed in Chapter 1, qualitative methods of risk assessment cannot necessarily be relied on to give correct or informative results in general, qualitative exposure assessments can be valid and useful under certain conditions. Specifically, suppose that the dose-response relation, representing the conditional probability of illness given exposure, increases from near zero to near one over a relatively narrow range of exposures – that is, a range of exposures that is narrow compared to the width of the overall frequency distribution for exposures. Then it is useful to divide the exposure axis into three contiguous ranges, which can be called "Low", "Medium", and "High", where "Low" refers to the exposures that give illness probabilities near zero, "High" refers to the exposures that give illness probabilities near 1, and "Medium" refers to exposures in between. As long as the transition from "Low" to "High" exposures occurs over such a narrow range that "Medium" exposures are unlikely and contribute little to risk, it matters little exactly where the boundaries among the three intervals are set. Fuzzy terms such as "Low" and "High" (possibly excluding the

"Medium" category altogether) for exposures then carry a useful natural meaning with respect to the probability of illness, i.e., a "Low" exposure is one that has little or negligible probability of causing illness, while a "High" exposure is one that has a substantial probability of causing illness.

The exact location of the boundary between "Low" and "High" (if only these two qualitative terms are used), or separating "Low", "Medium" and "High" (if all three are used), need not be specified precisely when risk is insensitive to these details. In this setting, quantitative risk is driven by the frequency of "High" exposures. In effect, the dose-response relation endows qualitative exposure values with a natural interpretation such that risk is proportional to "High" exposures. Indeed, Van Gerwen *et al.*, (1998, 2000) have suggest using simple but useful "Attention Values", such as those in Table 3, to help to decide quickly whether specific exposure levels to specific bacteria are likely to pose significant human health risks worthy of closer study and more accurate quantitative modeling. These values might be re-interpreted as rough estimates of the boundary between "Low" and "High" exposures for the different bacteria.

Table 3. Attention Values and Mortality Ratios for Several Bacteria

Bacteria	*AV* (cfu)	Mortality ratio (%) †‡
Bacillus cereus	1×10^5	0·0010
C. jejuni	500	
Clostridium botulinum type A	100	15
C. perfringens	1×10^5	0·10
Escherichia coli	1×10^4	0·30
L. monocytogenes	100	32
Salmonella spp.	1	0·21
V. parahaemolyticus	1×10^4	0
†(Fatal cases)/(total cases of illness or intoxication for the organism) × 100.		
‡If several values were given in literature, the worst case value was taken.		

Source: Adapted from a larger table by van Gerwen *et al.*, 2000

In summary, rapid qualitative exposure assessments, focusing on how frequently "High" exposures are likely to occur if different interventions are made, can serve a useful screening purpose when there is enough knowledge to separate "High" exposures (those that pose a significant increased risk of illness) from "Low" ones, for the exposed populations, with little error. Such qualitative exposure assessments also lead naturally to quantitative and semi-quantitative refinements, as illustrated in the following example.

Example: Quick Qualitative Exposure and Risk Assessment

Problem: Suppose that a ban on animal drug A will increase the frequency of "High" exposures to strains of bacteria B in food servings by 5% over its pre-intervention base level (i.e., new level = 1.05 × old level), while eliminating all A-resistant bacteria B in food servings. Susceptible cases of B infection take 6 days on average to resolve themselves, while resistant cases take 8 days. If, prior to the intervention, 80% of human cases of B infection are susceptible to A (with 20% being resistant), then what is the net impact of the intervention on days of illness?

Solution: Prior to the intervention, the average duration of an illness caused by B is:
Pr(susceptible) × E(duration | susceptible) + Pr(resistant) × E(duration | resistant)
= 80% × 6 + 20% × 8 = 4.8 + 1.6 = 6.4 days.
(E(duration | Y) denotes the conditional expected value of duration, given condition Y.) Following the intervention, the average duration of illness is 6 days, since almost all cases are now susceptible. The average frequency of illnesses due to contaminated food servings is 1.05 times greater than before, but the illness-days per illness has been reduced by (6/6.4) = 0.9375. Therefore, the net change in illness-days per year in the population is a reduction by a factor of (1.05) × (0.9375) = 0.9844. That is, illness-days per year decrease by 1.56% of the base level.

This example illustrates that quick estimates of the net effects of interventions can be developed by multiplying relevant factors, here defined using fuzzy but useful terms such as "High" exposure. Defining, measuring, or modeling the exposures more precisely would not necessarily improve the reliability, accuracy, or ease of interpretation of the predicted changes in exposures and risks. Therefore, the rough analysis just presented may be the most appropriate level of exposure assessment and risk assessment for these data (van Gerwen *et al.*, 1998.)

9. SUMMARY AND CONCLUSIONS

Exposure assessment models predict how human exposures to bacteria are changed by different risk management interventions. Both susceptible and resistant strains of bacteria are typically of interest for interventions that affect animal antibiotic use. Relatively well-developed techniques of microbial risk assessment (Haas *et al.*, 1999) can be applied to estimate the exposures to different bacterial strains in different foods under current conditions using epidemiological, microbiological (e.g., genotyping and other typing) and process-specific data. "What-if" predictive modeling is usually essential to predict the probable changes in human exposures that would be caused by different proposed interventions.

Three main types of predictive exposure models have been discussed. *Process simulation* models (including product-of-factors and farm-to-fork models) estimate the changes in microbial loads from point to point along a process leading from original sources (e.g., farm animals) to exposed targets (e.g., consumers or patients.) *Attribution-based* models begin with current cases per year and estimate the fraction that could be prevented by interventions that remove or block exposures from one or more sources. *Systems dynamics* models describe the high-level dynamic flows of bacteria between resistant and susceptible types in humans and animals under selection pressures from antibiotic use; and of people and animals between different health categories (e.g., ill and well, in the simplest cases.) Of these approaches, process simulation exposure models have been the best developed and most widely used in microbial risk assessment. Detailed illustrations for *Campylobacter* (Christensen *et al.*, 2001; Rosenquist *et al.*, 2003) and *Salmonella* (WHO/FAO, 2002) and other pathogens are available and have been used to illustrate several key ideas in this chapter.

However, despite considerable amounts of predictive microbiological information and human behaviour (e.g., hand-washing, kitchen hygiene, and cooking habit) data collected for such models, there are usually still important data gaps for details that affect final exposures, such as storage, preparation and cooking histories. This motivates development of alternative approaches that do not require unavailable details to provide useful exposure estimates. *Aggregate statistical exposure models* (e.g., a linear regression of historical bacterial illness rates against contemporaneous contaminated chicken carcasses produced) and statistical models that use prevalence of contamination (e.g., on processed carcasses or at retail) as a surrogate for exposure are methodologically flawed and should not be used for risk assessment. The *product-of-factors* and *attribution-based* approaches in this chapter attempt to replace the many low-level details of process simulation models with a smaller number of empirically estimated parameters, such as the ratios of microbial loads at successive measurement points in the production chain and the preventable fractions of exposure coming from different sources, respectively. *Systems dynamics* approaches may be necessary to model the emergence of resistance and to predict the time courses of increases or decreases in resistance among human and animal bacteria caused by interventions. However, as illustrated in this chapter, systems dynamics models can potentially make use of a relatively small number of high-level descriptive variables (e.g., ill animal, ill human, resistant animal, and resistant human fractions) that can be estimated from available data. Chapters 6-8 illustrate these approaches in greater detail.

Chapter 5

Dose-Response Modeling and Risk Characterization

This chapter completes the description of risk quantification by discussing dose-response modeling, health consequence modeling, and risk characterization, including uncertainty, variability, and sensitivity analyses. A principal goal is to show how to combine exposure information with dose-response and health consequence information to predict the probable human health consequences of exposure patterns created by different risk management decisions. Technical methods for displaying such risk information – and uncertainty about it – to others are also discussed. The focus in this chapter is on *technical* characterization of risks and uncertainties. The broader challenges of effective risk communication to different audiences and for different purposes, mentioned in Section 7 of Chapter 2, are not addressed.

A key insight is that *detailed quantitative exposure modeling and dose-response modeling are not always necessary* for microbial and antimicrobial risk assessment. It is often possible to replace both steps with simple approximate regression-type ("structural equation") models. These describe the expected number of adverse consequences (e.g., illnesses, illness-days, or QALYs lost) per year caused per serving ingested. This is the crucial information needed for risk characterization. It is not always possible (and may not be desirable) to create and validate separate exposure assessments and dose-response models to estimate risks, if simpler methods suffice to estimate risk-per-serving and how it changes for different risk management interventions.

1. INTRODUCTION TO DOSE-RESPONSE MODELING

Dose-response models quantify the conditional probability of illness caused by each level of exposure; thus, the term *exposure-response model* is also appropriate. In experimental studies, such as feeding trials that deliberately expose volunteers to known quantities of bacteria in selected food vehicles, "dose" is usually defined as the number of colony-forming units (CFUs) of bacteria ingested (or, more conveniently, as the common logarithm of this number, since the range of administered doses typically spans several orders of magnitude.). "Response" refers to development of illness caused by the ingested bacteria.

Naturally occurring exposures are often more complicated, perhaps involving repeated exposures via multiple servings of one or more foods over some time interval. Results from outbreak studies of naturally occurring food poisoning outbreaks are usually summarized by *attack rates* (what fraction of those exposed got sick?) and by an estimate of the order of magnitude of the exposures received (CFUs ingested) during the outbreak. Parametric statistical dose-response model curves can be fit to such data using dose-response modeling software or general-purpose commercial statistical analysis software, via maximum likelihood estimation (MLE) or other algorithms (e.g., Haas *et al.*, 1999; Cox 2001, Chapter 3).

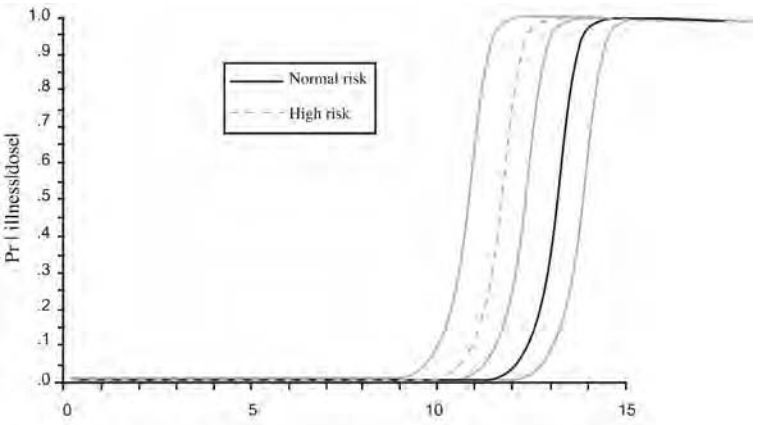
Figure 1 shows an example of a dose-response model developed for *Listeria monocytogenes* in ready-to-eat foods. A specific parametric dose-response model was assumed (the exponential model, discussed later) and fit to epidemiological data for immunocompromised ("High risk") and non-immunocompromised ("Normal") subpopulations. The dark solid curve in Figure 1 is the estimated dose-response model for the "Normal risk" subpopulation. The dashed line above and to the left of it is the dose-response model for the "High risk" subpopulation. The lighter gray curves indicate estimated statistical confidence bands around these best-estimate curves – an upper confidence band for each (corresponding to the upper end of the 95% confidence interval estimated for the parameter of the exponential dose-response model), and a lower 95% confidence band for the right-most (Normal) dose-response model.

As in Figure 1, it is often necessary to fit separate dose-response models to "normal" and "susceptible" subpopulations within the general population to account for inter-individual variability in dose-response relations. While more than two gradations of susceptibility can potentially be modeled using finite mixture distributions, distinguishing between only two levels or response "types" in the population, i.e., susceptible and normal, often suffices to explain most of the variability in the data.

Dose-response modeling may involve more than just quantifying the probability of a dichotomous outcome such as "illness". If different degrees

or severities of illness are possible, ranging from mild through severe to fatal (Buzby *et al.*, 1996), then a *health consequence model* describing the conditional probabilities of different levels or severities of health outcomes, given that illness occurs, is needed to augment the conditional probability of illness as a function of exposure. Section 3 of this chapter briefly discusses health consequence modeling. The appendix to Chapter 6 provides more detailed calculations for a case study example. In general, risk characterization requires describing the severities as well as the frequencies of adverse health outcomes caused by exposures.

Figure 1: Example Dose-Response Function for *Listeria monoctogenes*



Source: FAO/WHO, 2001. http://www.who.int/foodsafety/publications/micro/en/may2001.pdf

Note on terminology: "Hazard Characterization". In microbial risk assessment (MRA), dose-response modeling is often referred to as "hazard characterization". We will use the terms *dose-response modeling* and *exposure-response modeling*, as the term "hazard characterization" might better be applied to description and characterization of microbial hazards.

1.1 Definitions and Purposes of Dose-Response Modeling

Following a U.S. National Academy of Sciences framework for risk analysis (Jaykus, 1996), the U.S. FDA, CDC and USDA have described dose-response assessment as "The determination of the relationship between the magnitude of exposure and the magnitude and/or frequency of adverse effects" (www.foodsafety.gov/~dms/lmriskgl.html). The *Codex Alimentarius* Commission states that "For biological or physical agents, a dose-response assessment should be performed if the data are obtainable."

Dose-response modeling describes the probabilistic causal relation between exposures received and the frequency and severity of resulting adverse health consequences, including illness-days and death. For example, it quantifies the "exposures → illnesses" link (by specifying the conditional probabilities for illness given exposure) for each susceptibility type in the following causal graph model:

$$\begin{array}{c} \text{act} \rightarrow \text{exposures} \rightarrow \text{illnesses} \rightarrow \text{consequences} \\ \qquad\qquad\qquad\qquad\qquad\quad \uparrow \\ \qquad\qquad\qquad\qquad\quad \text{susceptibility} \leftarrow \text{covariates} \end{array}$$

As explained in Chapter 4, surrogates for ingested doses can be used, such as extent of contamination on meats at retail or at other points prior to the ingested servings. In this case, it is still possible to develop exposure-response models for this link, with the interpretation that the "illnesses" node describes the conditional probabilities of illnesses, given whatever has been observed. Bayesian network inference algorithms and techniques such as marginalizing out unobserved quantities, as mentioned in Chapter 4, can be used for this purpose.

It remains important to distinguish between changes in probabilities of illness and adverse health consequence *caused* by changes in exposures, on the one hand, and mere statistical *associations* between exposures and probabilities of adverse health consequences, on the other. Currently even authoritative regulatory and advisory bodies experienced in microbial risk assessment and/or antimicrobial risk assessment sometimes blur this crucial distinction.

Example: Association vs. Causation in Hazard Characterization

Current risk-related terminology and definitions are not always as clear and useful as they should be. For example, the *Codex Alimentarius* gives the following definition of hazard characterization: "The qualitative and/or quantitative evaluation of the nature of the adverse health effects associated with the hazard. For the purpose of Microbiological Risk Assessment the concerns relate to microorganisms and/or their toxins….The purpose of this step is to provide a qualitative or quantitative description of the severity and duration of adverse effects that may result from the ingestion of a microorganism or its toxin in food. A dose-response assessment should be performed if the data are obtainable." (www.foodriskclearinghouse.umd.edu/pversion/Codex_MRA.htm.) This description conflates the two very different concepts of association and causation. Evaluating the nature of the adverse health effects *associated with* a hazard such as a foodborne

bacterium is not the same as describing the adverse effects that may *result from* (i.e., be caused by) it.

As an example, infection with *E. faecium* bacteria can be associated with severe consequences, including death, even if the *E. faecium* infection itself does not cause these consequences. The reason is that *E. faecium* does not normally infect people with healthy immune systems. Thus, susceptibility to *E. faecium* infection may be a marker for a severe underlying health condition (compromised immune system) that, in turn, contributes to and is associated with various adverse health effects. In causal graph terms, confounding by severe illness, as in the diagram:

E. faecium infection ← *Compromised immune system* → *adverse consequences*

should not be conflated with characterization of effects caused by infection, diagrammed as the link *E. faecium infection* → *adverse consequences* in a model such as:

E. faecium infection → *adverse consequences* ← *Compromised immune system.*

In words, what we care about is the change in illness-days of different severities, deaths, QALYs lost, etc. incurred per year (or per capita-year) in a population that could be *prevented* by actions that reduce *E. faecium* exposures. This is generally not the same as (and will usually be smaller than) the adverse consequences that are statistically *associated* with *E. faecium* infections, as the latter include the effects due to confounding by compromised immune function.

A successful quantitative exposure-response model provides a mathematical function relating exposure levels (and possibly other covariates, such as membership in a susceptible subpopulation) to probabilities or expected frequencies of adverse consequences in exposed individuals. Uncertainty bands (Section 5.1) show uncertainties about consequence probabilities or rates at different exposure levels.

2. MICROBIAL DOSE-RESPONSE MODELING

This section summarizes several approaches to quantitative dose-response modeling suitable use with feeding trial and outbreak data.

2.1 Empirical Statistical Dose-Response Models

If experimental data are available, e.g., from feeding studies, then any parametric statistical model that adequately describes the dose-response data

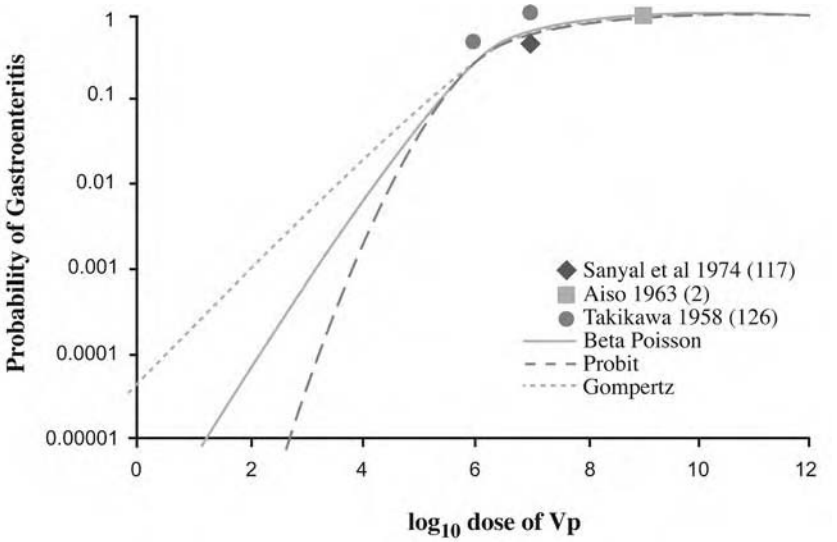
over the range of observations may be used to summarize it. The adequacy of a model-based description of data can be assessed by standard regression diagnostics (Haas *et al.*, 1999). Common parametric statistical risk models include logit, probit, log-normal, log-logistic, log-probit, Gompertz, and Weibull models. Within each family, the parameters of the curves can be fit to (i.e., estimated from) data via maximum-likelihood estimation (MLE) or other statistical curve-fitting algorithms available in commercial statistical and risk analysis software packages. (See WHO/FAO, 2003 and Cox, 2001, Chapter 3 for statistical methods for risk assessors). Nonparametric smoothing methods for fitting smooth curves to scatterplot data may also be used. Approximate uncertainty bands for empirical curves can be obtained by fitting curves within the selected parametric family to many different randomly selected subsets of the data using bootstrap or other resampling algorithms; see Haas *et al.*, 1999 and WHO/FAO, 2003 for relevant statistical methods of microbiological dose-response modeling.

However, in general, extrapolating empirical statistical models outside the range of the experimental data, especially to low doses, is *not* justified. Different parametric models fit to the same data can produce low-dose risk extrapolations that differ by orders of magnitude. For example, Holcomb *et al.* (1999) reported that "The statistical models proposed in the literature and chosen for comparison purposes were log-normal, log-logistic, exponential, Beta-Poisson and Weibull-Gamma. These were fit to four data sets also taken from published literature, *Shigella flexneri, Shigella dysenteriae, Campylobacter jejuni,* and *Salmonella typhosa*, using the method of maximum likelihood. … Within any given data set, the infectious dose estimated to affect one percent of the population ranged from one order of magnitude to as much as nine orders of magnitude, illustrating the differences in extrapolation of the dose response models." In light of such findings, empirical curve-fitting is best viewed as an *interpolation* approach for describing potentially large amounts of experimental data with a smaller number of parameters. Extrapolations outside the range of the data should be avoided.

Example: Dose-Response Extrapolations for *V. parahaemolyticus*

Figure 2 illustrates three different parametric statistical dose-response curves, fit by maximum likelihood estimation (MLE) to the same human feeding study data (from three studies, named in the figure and cited in FDA-CFSA, 2001). All three models pass through the experimental data points quite closely, but they give very different extrapolated predictions for risks at low doses. None of the three satisfactorily describes empirical risks estimated from epidemiological data (*ibid*).

Figure 2. MLE Beta-Poisson, Gompertz, and Probit Dose-Response Curves Fit to Pooled Human Feeding Study Data for *V. parahaemolyticus* (Vp).



Source: US FDA-CFSA, 2001. http://www.cfsan.fda.gov/~dms/vprisk5.html

As noted by the authors, "Consideration of the predicted density of pathogenic *V. parahaemolyticus,* the number of raw oyster servings for the Gulf Coast summer harvest and the likely number of illnesses occurring (CDC personal communication), strongly suggests that the predicted risks per serving based on dose-response curves shown in [Figure 2] are not plausible. Consequently, direct extrapolation of the dose-response under conditions of exposure in the feeding trials is not supported by the epidemiological data." Their suggested explanations illuminate the difference between purely statistical curve-fitting and modeling based on relevant biological and medical knowledge: "The human feeding trials were conducted under conditions of concurrent antacid administration. For *V. cholerae*, the ID_{50} [the dose at which 50% of subjects become infected] observed in feeding trials is known to be substantially lower when *V. cholerae* is ingested with antacid versus no antacid. The same effect is likely to be the case with *V. parahaemolyticus*. It is also possible that food matrix or immunological effects of preexposure to the organism, including antibodies/vaccines, contribute to the apparent difference in dose-response obtained under experimental versus natural conditions."

This example illustrates the critical importance of validating dose-response models (ideally, with data not used to develop them) before applying them to draw inferences about health risks in the real world.

In principle, empirical *multivariate dose-response functions* can be developed to predict illness risk as a function of not only dose variables

(such as the amounts ingested on successive occasions and the time intervals between them), but also of the age, sex, individual susceptibility and other covariates of the exposed subjects. Such functions describe the conditional probabilities of adverse effects (both frequency and severity) from exposures in different subpopulations, defined by different combinations of covariate values. Multivariate dose-response models quantify the following sub-diagram:

$$\text{dose} \rightarrow \text{illnesses} \leftarrow \text{covariates}$$

where "covariates" may refer to a vector of individual-specific factors that affect the relationship between dose variables and probability of illness. At present, univariate dose-response models are more commonly used in microbial risk assessment and antimicrobial risk assessment.

Technical Note: Multivariate dose-response modeling: To create a simple multivariate dose-response model, decision tree algorithms can be used to automatically bin more detailed dose-response data (giving the observed responses for many individuals exposed to different combinations of dose variables and perhaps other factors) into a few aggregate dose intervals (or combinations of intervals, if there are multiple risk factors in a multivariate dose-response model) that predict similar response levels (Zhang and Singer, 1999.) Such approximate dose-response relations can then be smoothed in a number of ways to construct simple, approximate smooth dose-response functions and confidence intervals directly from the more detailed data (Chaudhuri *et al.*, 1995, Loh, 2002). These techniques, while increasingly familiar and well-developed in biostatistics, are not yet commonly used in antimicrobial risk assessment.

2.2 Biologically Motivated Statistical Dose-Response Models

The probability that enough ingested organisms survive to reach a site where they initiate infection has been calculated in simplified biomathematical models of the probabilistic survival and infection processes (e.g., Teunis *et al.*, 1999). This approach leads to a catalog of parametric dose-response models appropriate for different simplifying assumptions about the disease process. They include the exponential, one-hit, multi-hit, Beta-Poisson, Weibull-Gamma, negative binomial, and threshold models, as well as mixture distribution models for populations with heterogeneous dose-response parameters. These models can be fit to experimental data by maximum-likelihood estimation (MLE) or other parametric statistical curve-fitting algorithms (Haas *et al.*, 1999). Unlike purely empirical models, these models provide a theoretical basis for extrapolating beyond the range of the

data used to fit them, at least to the extent that their underlying assumptions provide useful approximate descriptions of biological reality.

Example: The Exponential ("One-Hit") Dose-Response Model

If x colony-forming units (CFUs) of a pathogen are ingested and each independently causes illness with probability p, then the probability that no illness results is $(1 - p)^x$, and hence the probability of illness is $1 - (1 - p)^x$. For small p and large x, this is closely approximated by the exponential formula:

$$\Pr(\text{illness} \mid \text{dose} = x) = 1 - \exp(-p*x). \qquad \text{(Exponential Model)}$$

This formula also holds if x is the mean of a Poisson-distributed random dose (Haas *et al.*, 1999). Even for x = 100 and p = 0.001, the approximation is fairly good. The exact (binomial) probability of illness is $1 - (1 - p)^x = 0.09521$, while the exponential model approximation is $1 - \exp(-p*x) = 0.09516$. For realistic sizes of x (e.g., 10^{10}) and p, the approximation is even more accurate. The parameter p can be estimated (e.g., by MLE or regression) from data on the proportions of people who become sick at different dose levels. Figure 1 illustrates exponential dose-response curves fit to data on listeriosis rates in both normal and susceptible populations. In practice, there are important additional statistical considerations (e.g., fitting a dose-response curve to data in which the doses are uncertain requires that a model be used that allows for an uncertain x; while acknowledging that p may differ for different people leads to a mixture model in which the distribution of p values in the population must also be estimated.) However, the simple exponential model provides a reference model and point of departure for more sophisticated models. For some pathogens, including *L. monocytogenes*, the exponential model provides useful fits to many data sets.

The Beta-Poisson Model

One of the most important and widely used parametric dose response models in microbial risk assessment is the *Beta-Poisson model*. Conceptually, this model treats the effective dose (e.g., the number of infectious colony-forming units ingested among the random total number ingested, which is assumed to be Poisson-distributed) as a random variable with a binomial distribution. The success probability, p, of this binomial distribution is itself modelled as a Beta-distributed random variable. The probability of illness given measured exposure depends on the (random) effective exposure. For further motivation and explanation of the Beta-Poisson model, see Teunis *et al.*, 1999 and Teunis and Havelaar, 2000.

The exact Beta-Poisson model is commonly approximated by the parametric formula:

$$\Pr(\text{illness} \mid \text{Dose}) = 1 - [1 + (\text{Dose}/\beta)]^{-\alpha} \quad \text{(Approximate Beta-Poisson Model)}$$

The two parameters α and β, interpreted as shape (steepness) and location parameters, respectively, can be estimated from data on the proportion of subjects responding at different dose levels by an iterative MLE-type algorithm (Haas *et al.* 1999). Uncertainty bands for the fitted model can be estimated by resampling methods, e.g., from the variation of best-fitting curves across multiple randomly selected subsets of the data.

The middle dose-response curve in Figure 2 is a Beta-Poisson model, shown as one of the three parametric models fit to the available feeding study data. That the Beta-Poisson model can be interpreted as a mixture distribution model for random effective exposures makes it a natural model for a wide variety of settings.

Technical Note: Beta-Poisson Model. The probability of infection at a given dose can be calculated analytically if the probability that each ingested CFU initiates an infection has a beta-distribution, perhaps reflecting differences in virulence and/or susceptibility across microorganisms and hosts, respectively. (It is the Kummer confluent hypergeometric function.) The above approximate Beta-Poisson formula is accurate for α much smaller than β and for β much larger than 1. However, the statistical confidence bands (based on confidence intervals for the parameters α and β) in the approximate model can be quite inaccurate, creating an incentive to use the exact model (Teunis and Havelaar, 2000).

Example: Beta-Poisson Model for *V. parahaemolyticus* in Oysters

FDA-CFSA (2001) motivated the use of a Beta-Poisson dose-response model in its risk assessment of *V. parahaemolyticus* in oysters as follows:

> "It is likely that the density of pathogenic strains is spatially and temporally clustered in the environment to some degree. The average number of isolates that are pathogenic does not identify the extent of this clustering. To account for the probable spatial and temporal clustering of pathogenic strains relative to total *V. parahaemolyticus* densities, we have assumed a beta-binomial distribution for the number of pathogenic *V. parahaemolyticus* at the time of harvest. Under a beta-binomial distribution the percentage of total *V. parahaemolyticus* which are pathogenic varies from one sample of oysters (e.g. 12 oyster composite) to the next. Given the occurrence of outbreaks this appears to be a reasonable assumption but cannot be validated directly since extensive

quantitative surveys of pathogenic *V. parahaemolyticus* densities are not available. Specifically, based on the number of total *V. parahaemolyticus* (Vp_{total}), within a given composite, the number of pathogenic (Vp_{path}) present is assumed to be distributed as a binomial random variable with Vp_{total} trials (size parameter) and a probability of success (p) distributed as a beta random variable. The distribution of the probability parameter p is called a mixing distribution and the variation of this parameter across composites of oysters induces a clustering of pathogenic strains relative to total *V. parahaemolyticus*." (FDA-CFSA, 2001, http://www.cfsan.fda.gov/~dms/vprisk5.html.)

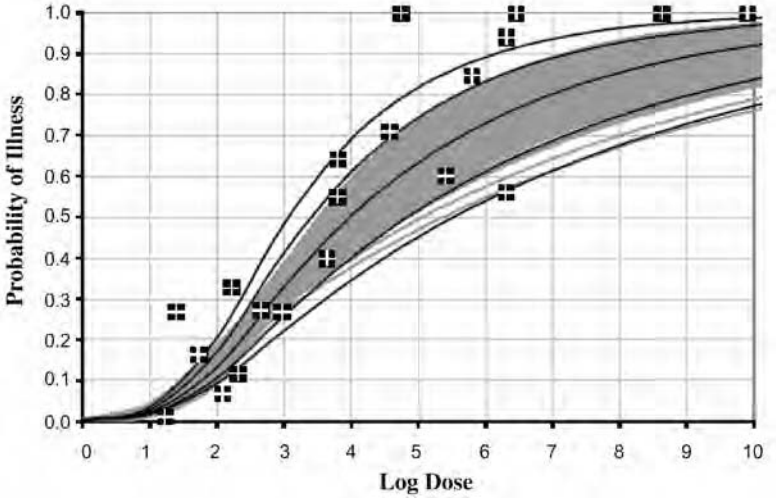
Thus, the Beta-Poisson model is here motivated by clustering of pathogenic strains.

Example: Uncertain Beta-Poisson Dose-Response Model for *Salmonella*

Figure 3 shows an estimated approximate Beta-Poisson dose-response model and uncertainty bounds for *Salmonella.* It was generated by treating observed data from outbreaks (i.e., estimated attack rates and exposures) as sample realizations of uncertain underlying outbreak processes. Multiple pseudo-data sets were constructed by sampling from estimated uncertainty distributions for the underlying outbreak data, and a Beta-Poisson dose-response model was fit to each. Figure 3 shows the range of resulting dose-response curves. As described by the authors, the figure "shows the comparison between the fitted curves and the expected value for the observed data. The upper bound, lower bound, expected value, 97.5th percentile and 2.5^{th} percentile for the dose-response curves fitted to the 5000 data sets are also shown. The fitted dose-response range captures the observed outbreak data quite well, especially at the lower and mid-dose range. The greater range at the high doses is due to the existence of several large-scale outbreaks at the lower- and mid-dose levels through which the curves attempt to pass, while the two high-dose data points are for relatively small-scale outbreaks that allow greater "elasticity" in the fit." (WHO/FAO, 2002)

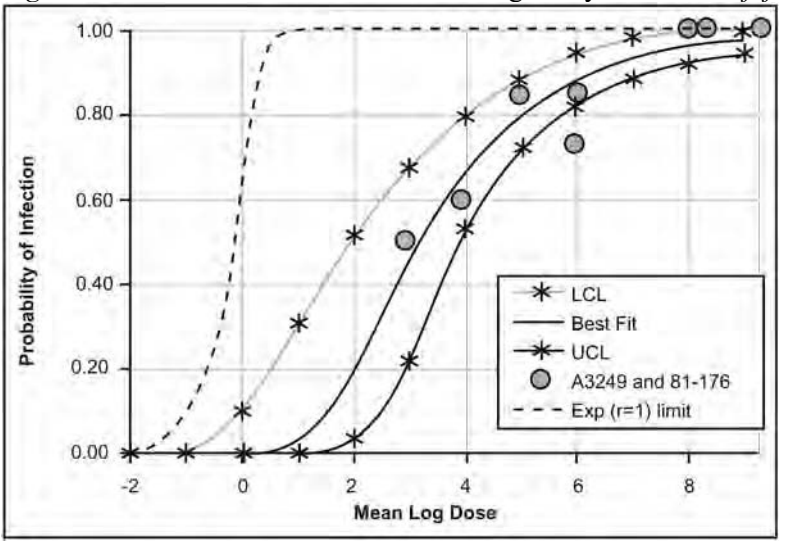
In addition to uncertainties in input data, e.g., true exposures and attack rates in outbreaks, values of parameters α and β may be different for different strains of a bacterium. For *Salmonella*, the decision to pool data across outbreaks was based on analysis of serovar-specific data suggesting that dose-response relations for the different serovars were similar (*ibid*).

Figure 3: An Uncertain Beta-Poisson Dose-Response Curve and Outbreak Data for *Salmonella*



Source: WHO/FAO, 2002

Figure 4: Beta-Poisson Curves and Feeding Study Data for *C. jejuni*



Source: FAO/WHO, 2002. ftp://ftp.fao.org/es/esn/food/campy.pdf

Example: Beta-Poisson Model for Campylobacter

Figure 4 shows upper and lower 95% statistical confidence limits (UCL and LCL, respectively) around the best-fitting exact Beta-Poisson model for the probability of infection, based on pooled data from human feeding experiments with two strains of *Campylobacter jejuni* (A3249 and

81-176). (The exponential limit, shown as a dashed curve in this figure, is an upper bound on the dose-response relation, derived by assuming a one-hit model with the probability of one cell initiating an infection of 1.) For *Campylobacter*, however, infection is not the same as illness, and the dose-response relation for illness is not as clear as that for infection. FAO/WHO (2002) described the issue and a proposed resolution for *C. jejuni* illness dose-response modeling as follows.

> "The probability of illness upon exposure to a dose of a pathogen is conditional upon the probability of infection. Stated another way, in order for an individual to become sick, the individual has to first become infected. The dose-response relationships described so far have estimated the probability of infection upon exposure to a dose. In order to estimate the probability of illness, the conditional probability of illness following infection is required. The human feeding trial data does not indicate a clear dose-response relationship for the conditional probability of illness following infection. For strain A3249, the data in the human feeding trials actually shows a decreasing trend for the conditional probability of illness with increasing dose. This observation has motivated some researchers (Teunis *et al.,* 1999) to postulate that perhaps upon exposure to a larger dose of some pathogens, the elicited host defenses are stronger, therefore reducing the probability of illness upon exposure to a very large dose compared to a moderate dose. The other alternatives that exist for the relationship of the conditional probability of illness following infection are that the probability increases with increasing dose or the probability is independent of dose. … In the case of the feeding trial data for *C. jejuni* A3249 the probability of illness decreases with increasing dose and as such a decreasing hazard function has been estimated (Teunis *et al.*1999). However, when the data for both strains are pooled the conditional probability of illness following infection does not exhibit a dose relationship but rather is randomly distributed …It may be appropriate in this case to use a dose independent ratio to estimate the conditional probability of illness. The conditional probability can be estimated from the feeding trial data. For A3249, out of 50 people that got infected at various doses, 11 got sick (22%), while for 81-176, out of 39 people that got infected at different doses, 18 got sick (46%). Overall, pooling all the data, a total of 29 people got sick out of 89 individuals that were infected (33%)."

The feeding study on which the data and dose-response model in Figure 4 are based has several limitations, including use of a milk vehicle rather than meat products of concern (e.g., chicken servings), and use of healthy young male volunteers rather than a more diverse and representative population. Section 2.4 discusses an epidemiological approach to estimating *Campylobacter* risks based on outbreak data.

2.3 Multi-Component Dose-Response Models

Dose-response relations can be decomposed into a sequence of components, e.g., representing internal dose received from a given external dose applied; probability of infection given internal dose; and probability of illness given infection. If these components can be estimated separately from available data, then the results can be composed (e.g., via Monte Carlo simulation of conditional distributions) to estimate the end-to-end exposure-illness dose response function. Separate estimation of components may help extrapolate results across species, if similarities and differences in relevant component processes are known. In practice, however, this decomposition strategy has usually been combined with simplifying assumptions about the components, leading to the biologically motivated parametric dose-response models (e.g., Beta-Poisson) already mentioned.

Use of a dose-response function estimated for one pathogen and host species as a surrogate for the unknown dose-response function of another pathogen and/or host species is not uncommon in past microbial risk assessments. Mechanistic models of species responses (e.g., based on differences in tissue pH and enzymology, retention times, volumes of food, etc.) can perhaps improve such extrapolations. However, such extrapolations across species are usually too uncertain to support confident risk assessment. They are not addressed further here.

2.4 Epidemiological Exposure-Response Models

An alternative to experimental data from feeding trials or animal studies is to look at the results of "natural experiments", i.e., situations in which multiple people have been exposed to high enough doses of a pathogen in some food or drink to become ill. For example, attack rates observed in some past outbreaks of campylobacteriosis are as follows:

- 47% following a business lunch (Brown *et al.*, 1988)
- 46.2% mean attack rate among 21 outbreaks of *Campylobacter* infection in England and Wales reported from 1992 to 1994. (Pebody *et al.*, 1997)
- 41% median attack rate among foodborne outbreaks of campylobacteriosis with at least six ill persons reported to the Centers for Disease Control and Prevention through the National Foodborne Surveillance Program in 1980-1982. (Finch and Blake, 1985)
- 19% to 67% in seven sequential outbreaks associated with a salad bar at one facility in Australia (Kirk *et al.*, 1997)
- 18.3% for a water-associated outbreak in Spain (Godoy *et al.*, 2002)

For the United States, an attack rate of about R = 41% among people getting very high doses (sufficient to cause an outbreak) may be a reasonable median value for foodborne cases of campylobacteriosis (Finch and Blake, 1985). How well this attack rate can be extrapolated to sporadic cases is unknown, but a simple assumption would be that in foods, as opposed to milk (used as a vehicle for administering *C. jejuni* in the feeding experiment in Figure 4), R = 0.41. This outbreak-based estimate of R = 0.41 is higher than the rate of 0.22 estimated for the A3249 strain from feeding study data described above, but is consistent with the rate of 0.46 for the 81-176 strain (FAO/WHO, 2002). The discrepancy between feeding study results and epidemiological data suggests that feeding study strains and subjects do not necessarily represent the strains and populations involved in outbreaks. Table 1 indicates the sizes of the adjustments that may be needed to extrapolate risk rates from healthy study populations to less healthy ones.

Another potential source of exposure-response information is epidemiological data on sporadic foodborne illnesses collected via case-control studies, cross-sectional surveys, or prospective cohort studies (relatively rare). A key statistical challenge for such data is to estimate exposures and conditional probabilities of adverse responses given exposures. Issues such as recall biases, omitted explanatory variables and confounders, uncertain model forms, and so forth, listed in Table 2 of Chapter 3, complicate valid statistical inference of dose-response functions from epidemiological data. However, advances in statistical methods, such as mixture distribution modeling with an unknown number of mixture components (Richardson and Green, 1997), increase the practicality of using epidemiological data for exposure-response modeling.

2.5 Practical Dose Response Models: Summary

Despite the range of theoretical approaches outlined above, in practice biologically motivated parametric dose-response models are the most common, and usually the best justified, models in widespread use. They are typically fit to data by a combination of MLE for point estimates and computationally intensive resampling techniques (e.g., bootstrapping algorithms) for confidence intervals and joint confidence regions for model parameters (Haas *et al.*, 1999, Chapter 7, c.f. p. 293).

Haas *et al.* (*ibid*, p. 98) state that "It has been possible to evaluate and compile a comprehensive database on microbial dose-response models." Chapter 9 of this monograph provides a compendium of dose-response data and dose-response curves, along with critical evaluations and results of validation studies, for the following: *Campylobacter jejuni* (based on human feeding study data), *Cryptosporidium parvum,* pathogenic *E. coli, E. coli*

O157:H7 (using Shigella species as a surrogate), *Giardia lamblia*, non-typhoid *Salmonella* (based on human feeding study data), *Salmonella typhosa*, *Shigella dystenteriae, S. flexneri,* etc., *Vibrio cholerae*, Adenovirus 4, Coxsackie viruses, Echovirus 12, Hepatitis A virus, Poliovirus I (minor), and rotavirus. Thus, for many food-borne and water-borne pathogens of interest, dose-response models and assessments of fit are already available.

It may be necessary to modify model parameters estimated from a limited subpopulation (e.g., healthy young male student volunteers) to predict risks in other populations. A simple adjustment is to multiply the model-predicted risks for the study population by *relative susceptibility factors* for other populations. Table 1 presents such factors for listeriosis in various populations. They were estimated from ratios of observed illness rates in these populations. On the other hand, dose-response models for several pathogens have been validated using outbreak data and other epidemiological sources, and the best-fitting models (often, the Beta-Poisson model) usually, but not always, fir available data quite well. Table 2 gives examples of pathogens for which exponential and/or Beta-Poisson dose-response models have been published, along with estimates of model parameters. (For the Exponential model, r is the probability that an ingested bacterium survives and causes infection. For the Beta-Poisson model, α and β are the parameters of the Beta distribution.)

Table 1: Relative Susceptibilities of Immunocompromised and Non-immunocompromised Populations to Listeriosis

France	**Relative Susceptibility Factor**
• Organ Transplant	2584
• AIDS	865
• Dialysis	476
• Cancer-Bladder	112
• Cancer-Gynaecological	66
• Elderly - over 65 years old	7.5
• Non-immunocompromised	1
United States	
• Elderly - over 60 years old	1.6
• Perinatal	839
• Non-immunocompromised	1

Source: http://www.who.int/foodsafety/publications/micro/en/may2001.pdf

Table 2: Dose-Response Parameters for Infection, and Mortality Ratios

Bacteria	Exponential (r)	Beta-Poisson (α/β)	Mortality ratio (%) †‡
C. jejuni A3249	3.52×10^{-6}	0.145/7.589	
L. monocytogenes	1.179×10^{-10}		32
Plesiomonas shigelloides	4.42×10^{-10}	0.057/1171	
Salmonella spp.		0.33/139.9	0.21 §
S. typhi		0.21/5531	6.0
S. typhi Quailes	2.14×10^{-8}	0.203/29173	
Shigella spp.		0.16/155	0.13
S. dysenteriae 1		0.5/100	
S. dysenteriae 1¶	2.05×10^{-4}	0.157/9·16	
V. cholera El Tor		2.7×10^{-5}/1.33	
V. cholerae 569b	1.76×10^{-9}	$0.508/7.52 \times 10^{7}$	

† (Fatal cases)/(total cases of illness or intoxication for the organism) × 100.
‡ If several values were given in literature, the worst case value was taken.
§ The average of deaths/cases for the organism of 1988 92 was used.
¶ Pooled dataset for strains A-1 and M 131.

Source: Adapted and abbreviated from van Gerwen *et al.*, 2000

2.6 Validating Dose-Response Models

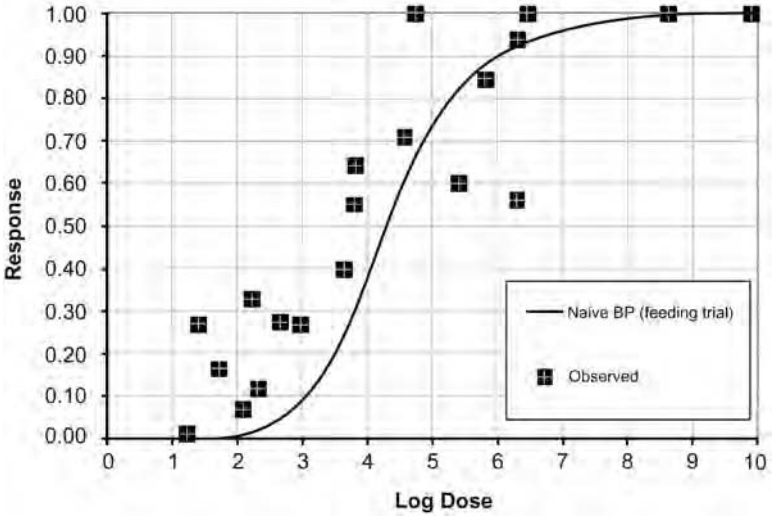
Dose-response models can and should be validated empirically, e.g., by analyzing epidemiological data from outbreaks (Haas *et al.* 1999, Chapters 7 and 9). The validation step allows the predictive accuracy of the models accuracy to be critically assessed.

One way to carry out dose-response model validation with outbreak data is to use the dose-response model, together with the estimated attack rates and durations of exposures and estimated quantities ingested, to predict the most likely dosages in the contaminated media that caused the outbreak, the most likely illness ratio during the outbreak, and levels of other observed quantities. The predicted levels can then be compared to actually measured or observed values recorded during the investigation of the outbreak *(ibid*). If the predicted levels do *not* match the empirically measured values (based on statistical tests of significant differences), then the exposure-response model should be corrected, e.g., by adding other relevant variables and/or by using the differences between predicted and observed values to select more appropriate mathematical model forms that reduce the differences.

Example: Best-Fitting Parametric Models May Not Fit Adequately

The following figure for *Salmonella* feeding trial data show that even the best-fitting model in a certain class of parametric models (here, the approximate Beta-Poisson dose-response family) may not adequately describe the observed data.

Figure 5: The Beta-Poisson Model Under-Predicts Low-Dose Risks



Source: WHO/FAO, 2002. (Naïve BP = approximate Binomial Poisson)

The parametric family of models is then said to be *misspecified* for the data, i.e., it is not appropriate for describing the empirical relation. In this example, the approximate Beta-Poisson model family is inappropriate for the data because even the best-fitting curve in the family dramatically under-predicts low-dose risks.

In a validated model, predicted values match validation data, as indicated by goodness-of-fit tests and model diagnostics such as plots of residuals. In this case, the dose-response model may be used to make predictions within the validated range of conditions. Remaining uncertainty in model parameters and predictions should be expressed through confidence intervals or uncertainty bands for single quantities (e.g., the mean illness rate in the population) and through joint confidence regions for multiple correlated quantities, such as the risks experienced by members of different subpopulations. Haas *et al.* (1999) provide details and examples. Unvalidated models should not be used for risk assessment.

2.7 Showing Uncertainty In Dose-Response Models

As illustrated in Figure 2 above and in Figures 3 and 4b of Chapter 2, realistic uncertainties about dose-response functions can create extremely wide ranges in predicted probabilities of illnesses for given doses. If the correct dose-response model is unknown and several models all provide adequate fits to the available data, then multiple plausible models may be used to carry out the rest of the assessment. In this case, the analysis can be organized and presented as a *model uncertainty decision tree* in which different modeling choices correspond to different branches in the tree. The results of the risk analysis at the end of each branch are contingent on the assumptions and modeling choices that lead to it. Different branches may be weighted by the relative strength of the evidence supporting them (Kang *et al.*, 2000). Bayesian model averaging (BMA) provides a version of this approach (Viallefont *et al.*, 2001; Keiding and Budtz-Jorgensen, 2004). The decision tree approach can also be used to present and analyze uncertainties due to choices of dose metrics, response definitions, and other modeling decisions, as well as choices of particular dose-response models.

Within a parametric dose-response model family, uncertainties about parameter values and model predictions can be quantified using statistical confidence bands and/or resampling techniques that compare results based on multiple subsets of the available data or based on multiple realizations of simulated data, as in Figure 3 above. Haas *et al.* (1999) and Cox (2001) discuss statistical methods for dose-response modeling, including cross-validation and bootstrap techniques for estimating the predictive power of a model and joint confidence regions for its uncertain parameter values.

Uncertainty about illness probabilities caused by a given dose is often dominated by uncertainty about the most appropriate dose-response model (sometimes called structural uncertainty), and a decision tree presentation of alternative modeling choices and the resulting predicted risks (or even a simple plot of different plausible dose-response curves, as in Figure 2) can express much of the relevant uncertainty with a minimal amount of statistical sophistication.

Statistical note: Weighting uncertain models. Branches in a model uncertainty decision tree can be weighted using formal statistical criteria for model selection such as the AIC, BIC, Mallows' criterion, cross-validation error, etc. (Cox, 2001). These measures are now built into many statistical software packages. In current risk analysis practice, subjectively judged weights of evidence are often used to combine results across multiple branches of the decision tree (Sielken *et al.*, 1995).

3. HEALTH CONSEQUENCE MODELING

In microbial risk assessment, the *number* of illness cases per year is often an adequate end point for risk assessment, as reducing it will also reduce all resulting adverse consequences such as illness-days and QALYs lost. But in antimicrobial risk assessment, including assessments of the human health impacts of animal antibiotic uses, it is important to quantify the incremental harm to patients caused by resistance to antibiotics. This corresponds to quantifying the following portion of Figure 1 in Chapter 3:

$$\Delta \text{ illnesses} \rightarrow \Delta \text{ consequences} \rightarrow \Delta\text{QALYs}$$
$$\uparrow$$
$$[\text{behavior, susceptibility, treatment}]$$

Here, "susceptibility" may be extended to include not only the susceptibility of the exposed individual to infection, but also the susceptibility of the infective bacteria to various treatments. (This diagram assumes that the consequences of illness are independent of the dose that caused it. If this is not true, then an arrow can be added from exposure to consequences.) The health consequences of an illness such as campylobacteriosis or salmonellosis may depend on the patient's behavior (e.g., care-seeking, compliance with physician instructions), susceptibility (e.g., immuno-compromised or not) and treatment (e.g., whether the physician prescribes an effective therapy.) Usually, these variables are marginalized out and only the conditional probability distribution for consequences (e.g., illness-days of different severities, early fatalities, or QALYs lost per illness) is quantified. Only if this distribution is significantly different for susceptible and resistant illnesses does resistance *per se* affect human health risk.

For example, if a risk management intervention eliminates *all* resistance to a certain antibiotic, A, in foodborne bacteria, but does not change the number of illnesses per year (because the resistant bacteria are replaced by susceptible ones) or their health consequences (e.g., because A is not used to treat illnesses and resistance to A does not affect responses to the antibiotics that are used), then the intervention will not cause any reduction in human health risk: the flow of adverse health consequences per unit time remains unchanged. On the other hand, if resistant bacteria cause an average of 2 excess illness-days per case compared to susceptible bacteria, then eliminating resistance will reduce illness-days per unit time, even if the number of illnesses per unit time remains fixed. Thus, to quantify the human health impacts of interventions that affect antibiotic use, it is essential to compare the health effects of resistant and susceptible cases.

Example: Consequence Calculations for Illness-Days

Problem: Suppose that 90% of illnesses caused by bacterium B are susceptible to antibiotic A, while the remaining 10% are resistant. The illness rate in a population due to bacterium B is 1E-4 illnesses per capita-year, i.e., one expected illness per 10,000 person-years at risk. The mean duration of illness is 6 days for susceptible bacteria and 8 days for resistant bacteria. The standard deviation of illness duration is 10 days in either case. How would the distribution of illness-days lost per year in this population change if resistance were eliminated?

Solution: The current expected number of illness-days per capita-year is:
Pr(illness occurs) × [E(duration | resistant) × Pr(resistant) + E(duration | susceptible) × Pr(susceptible)] = (0.0001) × (8 × 0.1 + 6 × 0.9) = 6.2E-4 days per capita-year.

(The probability of two or more illnesses in one person-year is small enough to be neglected.) If resistance were eliminated, this number would be reduced to: (0.0001) × (8 × 0 + 6 × 1) = 6.E-4 days per capita-year. In the population as a whole, summing over all individuals gives an approximate normal distribution for the number of illness-days per year. Its mean is: E(illnesses) × E(duration of an illness) = N × 1E-4 × 6.2 days before intervention and N × 1E-4 × 6 days after, where N = population size. Its variance is: E(illnesses) × Var(duration) + Var(illnesses) × E^2(duration) (Feller, 1968, p. 301) ≈ E(illnesses) × [Var(duration) + E^2(duration)] (since Var(illnesses) = Np(1 – p) ≈ Np = E(illnesses), where N = population size and p = 1E-4 = illness probability). Thus, Var(illness-days) = N × (1E-4) × 6.2 days × [100 + 38.44] = 0.0858 × N before intervention, and N × (1E-4) × 6 × [100 + 36] = 0.0816 × N after intervention. In summary, eliminating resistance reduces population risk from being an approximately normal distribution (if N is large) with mean 0.00062N and variance 0.0858N to a leftward-shifted, slightly narrower normal distribution with mean 0.006N and variance 0.0816N, where N is the population size.

As discussed in previous chapters, antimicrobial risk assessments that only quantify the number of cases per year in which resistance occurs, or in which it occurs and is attributed to an animal drug use and is treated with a resisted human drug, omit essential information. The key questions for risk assessment are: (a) Do these cases lead to increased human health harm; and (b) How would proposed interventions change the frequency (i.e., number of cases per year) and severity (e.g., illness-days or QALYs lost per case) of harm done? Question (a) should be addressed as part of hazard identification (see Chapter 2). It can be viewed as a screening question to determine whether the more detailed assessment needed to answer question (b) is likely to be worthwhile.

Example: Uncertain Human Health Hazards from Tylosin in Chickens

Macrolides such as tylosin are used in soluble and premix formulations in chickens to prevent and control several bacterial diseases and to promote health and growth. Such uses can potentially affect human health by changing microbial loads on chicken PRODUCTs and/or by selecting for macrolide-resistant pathogens and commensals reaching people on food commodities. Food-borne bacterial pathogens that are of specific concern as potential hazards to human health are *C. jejuni* and *C. coli,* both of which are found in live broilers (Wedderkopp *et al.*, 2003), chicken carcasses, and retail chicken products (Ge *et al.*, 2003; Musgrove *et al.*, 2003). Diagnosed cases of severe campylobacteriosis in humans may be treated with erythromycin or other macrolides. Empiric treatment of diarrhea with macrolides and fluoroquinolones is common, although these drugs do not necessarily have clinical benefits for most campylobacteriosis cases.

Although macrolides are important antibiotics in human medicine, this is relevant for risk assessment of tylosin use in chickens only to the extent that such use reduces the effectiveness of macrolides in human medicine. A human health risk exists only if there is potential to cause harm to human health, and not simply as a result of macrolides being important in human medicine. A clinical perspective on the treatment of *C. jejuni* infections is as follows:

> "Most *C. jejuni* infections are mild and self-limited; therefore, they do not usually require antibiotic therapy. Correction of electrolyte abnormalities and rehydration are usually sufficient. Treatment often is reserved for compromised hosts or persons with fever, increasing bloody diarrhea, or symptoms that last longer than 1 week. *C jejuni* is usually sensitive to erythromycin, gentamicin, tetracycline, ciprofloxacin, and clindamycin. Reports of erythromycin- and ciprofloxacin-resistant strains are increasing. In adults, placebo-controlled studies of erythromycin demonstrate no improvement in the clinical symptoms if given late in the course of illness but have resulted in decreased fecal shedding. If an appropriate antibiotic therapy was initiated within the first 4 days of illness, there was a reduction in the excretion of the organism; however, results regarding the resolution of symptoms were controversial. In contrast, early erythromycin treatment for children with bloody diarrhea shortened both the duration of diarrhea and excretion of microbes in the stool. Recommended duration for antibiotic treatment given for gastroenteritis is 5-7 days. Antimicrobial therapy for all bacteremic and immunocompromised patients with *C jejuni* should be selected based on a laboratory susceptibility test. Begin therapy with gentamicin, imipenem, third-generation cephalosporins, or chloramphenicol until susceptibility test results are available." (Ang and Nachman, 2003)

This perspective is consistent with clinical experience gathered over the past three decades. For example, for clinical effectiveness, despite some initial promising

reports on the efficacy of erythromycin in shortening the duration of *C. jejuni* campylobacteriosis (Nolan *et al.*, 1983), others soon found that "Although erythromycin significantly shortened the duration of *C jejuni* excretion, it appeared to exert no effect on the clinical course of the illness" (Robins-Browne *et al.* (1983); see also Anders *et al.*, (1982)). When investigators focused specifically on early treatment, they still found that that "Erythromycin rapidly eliminated *C. jejuni* from [human] stools…. Despite its bacteriologic effectiveness, erythromycin did not reduce the duration or severity of diarrhea, abdominal pain, or other symptoms" (Williams *et al.* 1989). In summary, macrolides are recommended as drugs of first choice for the small fraction of severe and high-risk campylobacteriosis cases for which antibiotic treatment is most warranted (and for which resistance screening is most likely). They have unclear clinical benefits in most other (non-severe) cases, and alternative treatment options are available for severe cases. It is not clear whether macrolide-resistant cases generally have worse clinical outcomes than macrolide-susceptible ones. (Some papers, such as Helms *et al.*, 2005, suggest that this may be the case, but without demonstrating whether resistance causes worse outcomes or is merely statistically associated with it following an incomplete statistical adjustment for comorbidity, e.g., in patients with other illnesses, such as AIDS, who may be at risk for both increased likelihood of resistance and increased likelihood of adverse health outcomes due to the other illnesses.) If not, then macrolide resistance *per se* would not present an incremental risk over and above the risks from campylobacteriosis. In such a situation, where hazard identification does not indicate a clear human health risk from resistance, it is often easiest to make progress by *assuming* that such a risk exists (e.g., that resistance creates an average of at most 2 excess days per of illness per case) and then carrying out the rest of the risk assessment, including human health consequence assessment, contingent upon this assumption. Chapters 6 and 8 illustrate this strategy in detail.

If hazard identification does identify some adverse consequences of resistance, such as increased treatment failure rates or prolonged duration of illness, then these consequences should be included in the risk assessment. This can be done by quantifying the average severity (e.g., illness-days or QALYs lost) for resistant and susceptible illnesses and estimating the change in the number of each that will occur if different risk management interventions are undertaken, as illustrated in the simple example calculation at the beginning of this section. Calculations with realistic data are illustrated in Chapters 6 and 8, using resistance to streptogramins as a case study.

4. RISK CHARACTERIZATION

Risk characterization provides the ultimate output of a risk assessment. It integrates hazard identification, exposure assessment, and dose-response information to determine the probable frequency and severity of adverse health effects in a population that are caused by exposure to a hazard. Characterizing the risks for different risk management interventions helps decision-makers choose among them. To further support practical risk management decision-making, risk characterization also includes characterization of current *uncertainty about risk*. This allows the value of gathering additional information to be assessed as part of risk management deliberation and decision-making, based on the potential value of such information (VoI) in enabling risk managers to make choices that are more likely to result in desired consequences (Yokota and Thompson, 2004).

4.1 Definition and Purposes of Risk Characterization

Different national and international bodies have attempted to describe the integration step in risk characterization non-mathematically, with varying degrees of success. For example, a Joint WHO/FAO Expert Consultation defined risk characterization as the "integration of hazard identification, hazard characterization and exposure assessment into an estimation of the adverse effects likely to occur in a given population, including attendant uncertainties". This give a useful impression of what is intended, but does not specify how the estimation is to be performed or presented, or define exactly what is meant by “adverse effects likely to occur in a population”. (*How* likely, in *how much* of the population, over what *time interval*, under what assumed *conditions*?) The US FDA has used this definition in applied work (http://www.foodsafety.gov/~dms/lmriskgl.html).

A more confusing exposition, developed as part of international microbial risk assessment (MRA) efforts, is:

> “Risk Characterization - The process of determining the qualitative and/or quantitative estimation, including attendant uncertainties, of the probability of occurrence and severity of known or potential adverse health effects in a given population based on hazard identification, hazard characterization and exposure assessment. … Risk Characterization represents the integration of the Hazard Identification, Hazard Characterization, and Exposure Assessment determinations previously described into qualitative or quantitative estimates of the likelihood and severity of the adverse effects which could occur in a given population, including a description of the uncertainties associated with these estimates.” (*Source:* http://www.foodriskclearinghouse.umd.edu/pversion/Codex_MRA.htm)

This does not specify what "The process of determining the qualitative and/or quantitative estimation" should include. (Would "I'll bet the risk is too high" or "I guess it's probably ok" be acceptable examples of "qualitative and/or quantitative estimation, including attendant uncertainties"? Why or why not?) In short, such verbal definitions and discussions do not make competely clear what is intended and acceptable for risk characterization. From this standpoint, more quantitative definitions such as "The change in the expected number of illnesses per year in a population that would be caused by implementing each risk management intervention" may be easier to understand and implement.

In summary, risk characterization is intended to show the predicted probable frequency and severity of adverse human health consequences (and other adverse effects of concern) for different risk management decisions. It typically presents expected impacts and confidence intervals for the number and severity of adverse outcomes per capita and per unit time. (With modern software, full conditional probability distributions can also be calculated and displayed.) Thus, *risk characterization relates decisions to their probable consequences in order to guide and inform improved risk-management decision-making.*

4.2 Desired Outputs of Risk Characterization

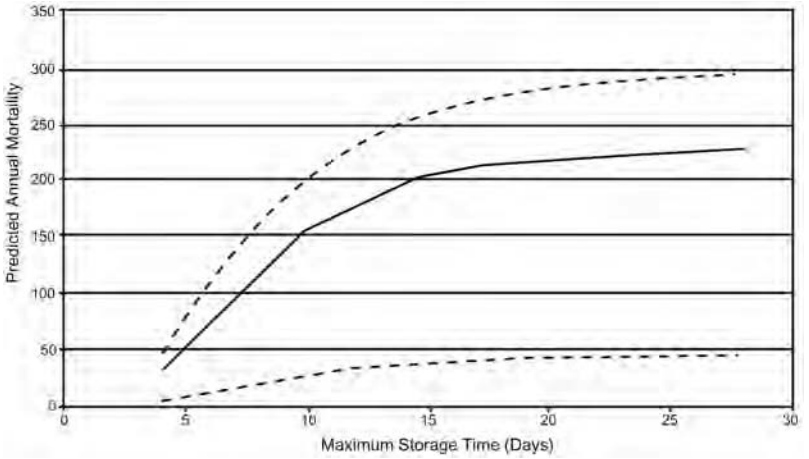
A well done risk characterization describes the microorganisms and/or resistance determinants of concern; the adverse human health effects that they cause (based on the hazard identification step); the frequency distribution of exposures and resulting illnesses in the population, with confidence limits, for different decisions; and expected values with confidence limits (or entire conditional probability distribution) for the frequency and severity of adverse effects for different risk management decisions (Haas *et al.*, 1999). Risk characterization outputs should include:

- *Expected risk metrics* (i.e., expected numbers of infections, illnesses of specified severity levels, mortalities, treatment failures, etc.) per year and per lifetime for a randomly selected member of the population;
- *Confidence intervals* around the expected risk;
- Expected risks and confidence intervals for members of identified *sensitive subpopulations* and *highly exposed subpopulations*;
- Expected numbers and confidence intervals for *total infections and illnesses with different levels of severity,* per year, per capita-year, and per capita-lifetime in the total population and identified subpopulations.

These individual risk and population risk metrics should be provided for each risk management decision being considered. The characterization should include any adverse effects that an intervention might cause as well as those that it might prevent, to provide decision-makers with a full accounting of the total change in human health from each intervention.

Example: Risk Characterization Outputs

Figure 6: Predicted Annual Mortality in the Elderly Subpopulation Attributable to Deli Meats as a Function of Maximum Storage Time



Source: FAO/WHO, 2001, http://www.cfsan.fda.gov/~dms/lmr2-6.html

Figure 6 shows one of the outputs from a risk assessment of *Listeria monocytogenes* (FAO/WHO, 2001). The solid curve shows the median estimate of the mortalities per year caused among the elderly subpopulation by *L. moncytogenes* in deli meats, for different maximum allowed storage times. The dotted curves represent the 5th and 95th percentiles of the uncertainty distribution (as assessed by Monte Carlo uncertainty analysis, discussed below.) This display shows how predicted risks in this subpopulation vary with the effects of different potential interventions that would limit the maximum storage times allowed for deli meats. Similar curves can be shown for the effects of such interventions for other foods or groups of foods (e.g., dairy products, produce, sea food products, etc.) and for other subpopulations and the U.S. population as a whole.

4.3 Methods for Risk Characterization

Basic Formulas

Conceptually, risk characterization is a purely arithmetic process that combines the results of exposure assessment and dose-response modeling to predict the overall probability of illness. For example, suppose that exposure assessment gives the probability of ingesting x colony-forming units of a pathogenic bacterium in a serving of a food commodity as some number, p(x) = Pr(dose in a serving = x), while dose-response modeling gives the conditional probability of illness from ingesting x CFUs as r(x) = Pr(illness | dose = x). Then the overall probability of illness from a serving is:

$$\text{Illness probability per serving} = \Sigma_x \Pr(x)*\Pr(\text{illness} \mid x) = \Sigma_x p(x)r(x),$$

where the sum is over all exposure levels. [Thus, in the terminology of Chapter 3, one "marginalizes out" exposure to get the *unconditional* probability of illness from a serving. This assumes that any longer-term effects that affect illness probability given dose, such as acquired immunity from frequent exposures to low levels of bacteria in this food product, are already taken into account in definig r(x).] If exposure is modeled as a continuous variable, then the above sum is replaced by an integral: illness probability per serving = $\int_x p(x)r(x)dx$. Thus, the core risk characterization calculation literally "integrates" exposure assessment and dose-response information. If the number of servings per year at dose level x is denoted by n(x) (= M*p(x), where M is the total number of servings per year) and the average consequence severity (e.g., illness-days or QALYs lost) per illness is c(x), then the population risk (i.e., expected number of illnesses per year and adverse consequences per year from exposure) can be expressed as:

$$\text{Expected illnesses per year in population} = \Sigma_x r(x)n(x)$$
$$\text{Expected adverse consequences per year in population} = \Sigma_x r(x)n(x)c(x).$$

(For continuous x, the sums are replaced by integrals.) Usually, c(x) is treated as a constant independent of x, i.e., the consequences of an illness do not depend on the dose that caused it, but this need not be assumed.

Example: Risk Characterization Calculations

Setting: Table 3, adapted from a risk assessment for *Listeria monocytogenes* (FAO / WHO, 2001), shows the estimated number of servings per year of ready-to-eat foods

carrying different levels of contamination (i.e., doses), as well as the estimated risk (i.e., probability of illness) per serving for each dose level.

Table 3: Baseline Number of Cases Predicted By a Dose-Response Model

Log of dose, x, at ingestion (Log_{10} CFUs per serving)	Average risk per serving at dose x = r(x)	Number of servings per yr at dose x = n(x)	Expected cases/yr. at dose level x = n(x) × r(x)
-1.5	1.69E-13	5.93 E10 (92.5%)	0.01
-.5	2.00E-12	2.50 E9 (3.9%)	0.005
.5	1.64E-11	1.22 E9 (1.9%)	0.02
1.5	1.71E-10	5.84 E8 (0.9%)	0.1
2.5	1.80E-09	2.78 E8	0.5
3.5	1.82E-08	1.32 E8	2.4
4.5	1.85E-07	6.23 E7	11.5 (0.6%)
5.5	1.85E-06	2.94 E7	54.4 (2.86%)
6.5	1.85E-06	1.39 E7	25.7 (1.35%)
7	5.88E-05	3.88 E6	228 (12.0%)
7.5	5.92E-04	2.67 E6 (0.0052%)	1580 (83.0%)
Totals		6.41 E10	1902.7

Source: Adapted from FAO/WHO, 2001. (Total cases were shown in the original as 2130, but is revised to 1902.7 here to be consistent with the other numbers.)

In this table, the column "Average risk per serving" at different dose levels summarizes dose-response information. For example, the estimated probability of illness from a serving containing ten million (10^7) CFUs of *L. monocytogenes* is shown as 5.88E-05, based on the dose-response model (see Figure 1). The column "Number of servings per year at the specified dose" summarizes the results of the exposure assessment. (The percentages of total servings are shown for the largest and smallest dose categories and several others; thus, 92.5% of servings are in the lowest category, where the average dose is only $10^{-1.5}$ = 0.032 cells per serving.) The final column gives the number of cases per year expected from servings at each dose level (and selected percentages). For example, 83% of all cases per year are estimated to come from the highest-dose servings, which account for only 0.0052% of all servings. (This illustrates the importance of not using prevalence as an exposure metric, mentioned in Chapter 4. Eliminating only the most-contaminated servings would prevent about 83% of all cases, while reducing the prevalence of contaminated servings by less than 0.01%, a statistically undetectable change.) The total population risk for number of illnesses per year is found by summing the final column (corresponding to the formula $\Sigma_x r(x)n(x)$), giving 1902.7 expected illnesses per year.

Problem: In Table 3, what would be the effect on population risk of an intervention that reduces all doses by a factor of 10? (For simplicity, assume that the numbers in Table 3 are correct, i.e., ignore model uncertainty for purposes of this example.)

Solution: Table 3a summarizes the revised numbers. (Rows not shown do not contribute significantly to the total.) The dose-response column r(x) is the same as in Table 3. The n(x) column is shifted up to reflect a 10-fold reduction in doses (i.e., microbial loads). Expected cases per year falls from 1902.7 to 44.65, i.e., reducing all microbial loads to 10% of their previous values reduces population risk to (44/65/1902.7) = 2.35% of its previous level. This reflects the disproportionate importance of the highest dose levels.

Table 3a: Baseline Number of Cases Predicted By a Dose-Response Model

Log of dose, x,	r(x)	n(x)	Expected cases/yr., n(x) × r(x)
2.5	1.80E-09	1.32E8	0.24
3.5	1.82E-08	6.23E7	1.13
4.5	1.85E-07	2.94E7	5.44
5.5	1.85E-06	1.39E7	25.72
6	1.85E-06	3.88E6	7.18
6.5	1.85E-06	2.67E6 (0.0052%)	4.94
Totals		6.41E10	44.65

In addition to expected values, risk characterization can produce entire probability distributions for outputs. These may be stated as probability distributions of the number of adverse human health consequences per year, if interventions only affect the steady-state risk levels created by a process. They may also show changes in consequences over time (e.g., to reflect changes in food consumption or preparation habits or in prescription practices and availability of new drugs), where applicable. The case study in Chapter 6 illustrates estimation of impacts over time.

In general, risk characterization may be viewed as calculating the unconditional probability distributions of outputs such as illnesses, adverse consequences (illness-days and early fatalities), and QALYs lost per year in an influence diagram such as the following:

$$\begin{array}{ccccccc} \text{act} & \rightarrow & \text{exposure} & \rightarrow & \text{illnesses} & \rightarrow & \text{consequences} \rightarrow \text{QALYs} \\ & & \uparrow & & \uparrow & & \uparrow \\ & & [\text{behavior} & & \text{susceptibility} & & \text{treatment}] = \text{type of case} \end{array}$$

[This is essentially Figure 1 of Chapter 3. The deltas indicating changes are left out here to emphasize that we can directly compare the probability

distributions for outcomes (e.g., QALYs lost per year) from *each* act in the set of alternatives being considered (including the *status quo*), rather than expressing them as changes relative to the *status quo* distribution of outcomes. While these perspectives are mathematically equivalent, expressing results as changes relative to the *status quo* may affect decision-makers differently from expressing the same information as total probabilities of outcomes (Plous, 1993).]

Given the results from:

(a) Exposure assessment (i.e., the conditional probability distribution of exposures, for each act);
(b) Exposure-response or dose-response modeling (i.e., the conditional probability of illness for each exposure pattern); and
(c) Consequence modeling (i.e., the conditional probability distribution of adverse consequences given illness, if consequences are conditionally independent of exposure given that illness occurs; and otherwise the conditional probability distribution of adverse consequences given illness and exposures)

the risk characterization step calculates, for each act being assessed, the resulting probability distributions for adverse consequences. As mentioned in Chapter 3, the probability distribution for outputs such as QALYs lost per year if a given act is chosen can be calculated by applying Monte Carlo uncertainty analysis (Thompson *et al.*, 1992) to the above influence diagram.

Monte Carlo Simulation-Based Risk Characterization

Monte Carlo-based risk analysis uses conditional probabilities to combine dose-response information (which may depend on individual covariates such as membership in susceptible subpopulations, as in Table 1) with exposure information (which may also depend on individual covariates) to estimate resulting risks in the population, and in subpopulations defined by combinations of values for the covariates. For example, the probability distribution for the number of occurrences of a specific effect, y, in a population over some period of time can be found as follows.

1. Create a random sample of simulated "individuals". If z denotes the vector of individual attributes and covariates, such as age, ethnicity, cooking habits, susceptibility class, etc., with joint probability density Pr(z), then "individuals" in the simulation are represented by realizations z(1), z(2), …, z(S) randomly sampled from Pr(z). Here, z(j) denotes the vector describing individual j, for j = 1, 2, …, S, where S is the size of the simulated population.

2. Assign exposure histories to individuals. Let Pr(x | z(j), d) denote the conditional probability of exposure x on day d, given individual covariates z(j). The probabilities of different exposure levels can be conditioned on any available information about the individual, z(j), and about day d, such as the season of the year. The conditional probability model Pr(x | z(j), d) is given by the output of the exposure modeling step. Let x(j, d) denote a randomly sampled realization drawn from the distribution Pr(x | z(j), d).
3. Assign individual dose-response functions, r(x | z(j)) = probability of effect y for an individual with description z(j) exposed to x.
4. Assign health outcomes to individuals and days. This is done by assigning an occurrence indicator value of 1 for outcome y in individual j on day d with probability r(x(j, d) | z(j)), and assigning an occurrence indicator value of 0 with probability 1 – r(x(j, d) | z(j)).
5. Sum the occurrence indicator values over all j and d (i.e., over all person-days in the population) to obtain the total simulated number of occurrences of y in the population over the time interval of interest.
6. Iterate steps 1-5 many times to simulate the frequency distribution for the number of occurrences of y in the population over the time interval of interest.

These steps provide a basic (not very efficient) simulation of the sum of individual risks on different days, with uncertainties about individual exposures x and descriptions z being marginalized out. The efficiency of the simulation can be greatly improved (e.g., by not returning all the way to step 1 on each iteration), but this simple scheme conveys the main ideas.

Repeated sampling of individuals (from the joint frequency distribution of individual covariates) and of exposures and dose-response relations given individual characteristics, allows all of the output risk metrics, confidence intervals, and confidence regions to be automatically calculated as accurately as desired. In practice, commercial Monte Carlo uncertainty analysis software products, such as Analytica™, @RISK™ or Crystal Ball™ automatically perform the required simulations, collect the results, and display the output risk metrics and uncertainty intervals, making Monte-Carlo based risk assessment practical even for practitioners with little understanding of underlying risk analysis principles (e.g., Vose, 2000). These software tools incorporate more advanced simulation techniques (such as antithetic variates for variance reduction, importance sampling, Latin Hypercube sampling, etc.) that increase the efficiency of the simulation process, greatly reducing the CPU-time needed to obtain accurate answers. In special cases, risk characterization calculations can also be carried out symbolically or analytically. However, the current state of practice generally

relies on Monte Carlo uncertainty analysis to obtain fast, accurate numerical answers. Burmaster and Anderson (1994) provide guidance and principles for using and documenting Monte Carlo uncertainty analysis in risk characterization.

Example: *Listeria monocytogenes* Risk Assessment Simulation Model

Table 4 summarizes the inputs and calculations performed in a Monte Carlo simulation-based risk assessment of the public health risks from *L. monocytogenes* (FDA-CFSAN, 2003) in twenty-three ready-to-eat foods in the United States. The risk assessment was used to characterize several risks, including the average risk-per-serving of listeriosis for each type of food. (Risks in elderly and compromised subpopulations and changes in risks from interventions of different magnitudes are also presented; see http://www.cfsan.fda.gov/~dms/lmr2-6.html.)

Although some of these steps (e.g., extrapolation from mouse dose-response data to human dose-response via an adjustment based on CDC population-level data) are of uncertain validity, as discussed in greater detail in the FDA-CFSAN report (see e.g., Appendix 11, http://www.cfsan.fda.gov/~dms/lmr2-toc.html), the steps in Table 4 indicate the range of inputs and conditional probability relations used in this Monte Carlo-based microbial risk assessment. Clearly, the approach requires a great deal of information. This motivates the search for simpler approaches that give equally valid answers.

Until they are carefully validated, simulation models and their results should be treated with healthy scepticism, especially if they make counterintuitive predictions – potentially, the most valuable and informative aspect of modeling when the results of the model are trustworthy. Potential weaknesses include:

- *Model validity:* If the simulation model logic is incorrect, the model may produce incorrect answers. For example, a simulation model may attribute large impacts to interventions that turn out to have small or no impacts in reality if the model falsely assumes that most exposures come from the sources identified in the model (e.g., consumption of food animal products) but real exposures come from other, currently unidentified sources (e.g., from flies and surface water.)
- *Input uncertainties:* Is the assumed dose-response model realistic enough to be useful? Is the assumed fraction of cases attributed to servings of the food commodity of concern approximately correct? Such input uncertainties are often hidden in the final output curves, which

Table 4: Steps in Simulation-Based Risk Assessment

1.	Distributions for contamination at retail for each food category.
2.	Distributions for the reference growth rate at 5°C for each food category.
3.	A distribution of home refrigerator temperatures in the United States, this distribution was used for all food categories.
4.	Distributions for post-retail storage time for each food category.
5.	The growth model used for all food categories. The growth model was triggered only for servings with one or more bacterium. This section calculated the exponential growth rate for the refrigeration temperature and multiplied that by the storage time.
6.	The maximum concentration for each food category. Post growth *L. monocytogenes* concentrations were truncated at this level. The maximum growth was temperature dependent with more growth allowed at higher refrigeration temperatures.
7.	A model representing the effect of reheating frankfurters on *L. monocytogenes* concentration, used for frankfurters only.
8.	Net contamination at time of consumption. Calculated with inputs from steps 1, 6, 7.
9.	Distributions of serving size for each food category.
10.	Distributions of dose at consumption for each food category. This is the final output of the 2D simulation. After collapsing the variability dimension to half-log dose bins, the output for each food category was conveyed to the 1D dose-response simulation for each population group.
11.	A distribution for variability of *L. monocytogenes* strain virulence in mice, with the implicit assumption that a similar range will be observed in humans.
12.	A distribution adjusting for variability in host susceptibility among humans, with three (High, Medium, Low) separate adjustments applied to represent different possible ranges. The adjustment increased the range of effective doses.
13.	The sum of strain variability (step 11) and host susceptibility distributions (step 12) obtained by 2D Monte-Carlo, with 100,000 variability iterations and 300 uncertainty iterations. The variability dimension was then collapsed to half log dose bins.
14.	Summation of the exposure assessment (step 10) and adjustment factor (step 13) for each food category
15.	The annual number of meals consumed for each food category.
16.	Addition of the dose-response adjustment factor that is applied in order to make the predictions consistent with CDC estimates of the annual death rate attributable to the population group (i.e., the median value in step 22).
17.	An intermediate calculation of the number of annual servings falling in each dose bin for each food category. This was obtained by multiplying the number of servings (step 15) by the fraction falling in each effective dose bin (step 14).
18.	Calculation of the death rate per serving for each dose bin (from step 14), using the dose-response function derived from mouse data.
19.	An intermediate calculation of the number of annual deaths for each dose bin and food category. This was obtained by multiplying the death rate per serving (step 18) by the number of servings for the dose bin (step 17).
20.	Calculation of the death rate per serving for each food category by summing across dose bins. This was obtained by summing the product of the death rate (step 18) and serving fraction (step 14) across all bins.
21.	Calculation of the annual number of deaths for each food category by summing across dose bins (step 19).
22.	Calculation of the total number of deaths by summing across food categories

Source: FDA-CFSAN, 2003. http://vm.cfsan.fda.gov/~dms/lmr2-a3.html

show *conditional* distributions of outputs for specific inputs and assumptions. Sensitivity analysis, discussed below, can help to demonstrate and build confidence in the robustness of the main conclusions to some key uncertainties (if the model logic is correct), such as details of the assumed dose-response function. But if these analyses demonstrate that conclusions are sensitive to some uncertain inputs, then results must be regarded as *contingent* on these inputs.

- *Representing interventions:* Effects of risk management intervention scenarios are often modeled as shifts in the probability distributions of microbial loads at processing, with the extent of the shifts being estimated from limited experimental trials. Without implementing the interventions at multiple plants and monitoring the results, there is no assurance that the true effects are accurately modeled.
- *Statistical modeling of parameter uncertainties:* Weibull, uniform, binomial, triangular, normal, log-normal, and other simple parametric uncertainty distributions for inputs predominate in most Monte Carlo models. They are usually selected because they are easy to fit to available data. Other distributions could be used instead. Although the choice of distributions to represent parameter uncertainties reflects subjective judgment and is somewhat imprecise, the final conclusions may not be very sensitive to these choices, provided that the selected distributions are flexible enough to fit available data well. For example, provided that means and standard deviations are estimated approximately correctly, the distribution of final microbial loads in many models will be approximately log-normally distributed (Druzdzel, 1994), regardless of the specific distributions used to model individual parameter uncertainties.
- *Complexity and verifiability of model results.* Simulation results based on Monte Carlo analysis cannot easily be independently verified, except by replicating the simulation. By contrast, simple back-of-the-envelope calculations are easier to check and make potential sources of disagreement easier to isolate for discussion. Such transparency may be essential to build shared understanding and convincing conclusions.

A challenge for any complex model that produces unexpected results is how to convince an open-minded sceptic that the results are correct. Merely exhibiting the assumptions and data for the model may not meet the needs of someone whose prior knowledge and experience make it difficult to believe the findings. We turn next to simpler quantitative risk assessment approaches that sacrifice the flexibility of uncertainty and variability modeling possible in simulation models in favor of increased transparency and clearer derivation of key conclusions from data.

4.4 Methods of Simplified Risk Characterization to Avoid

Rather than building up risk estimates from estimated dose-response functions and frequency distributions of exposures, it has sometimes been proposed that total risk in a population can be estimated from data and that aggregate (population-level) regression models or equations can be used to attribute shares of this total to particular sources. Thus, for risks of resistant cases of bacterial illness in humans, a simple statistical model such as:

Population Risk = Background + K*(Exposure from animal use)

could be fit to population-level data on the number of cases per year (i.e., "Population Risk") and the estimated "Exposure from animal use". Changes in "Exposure from animal use" would then be assumed to affect Population Risk through the coefficient K.

In general, such regression-based, attribution-based approaches may produce results that are causally meaningless (Freedman, 2004), as regression models (and population attributable risk fractions) do not necessarily correctly indicate the *changes* in effects (e.g., in Population Risk) that would be caused by changes in exposures or other inputs. (For example, if men have more exposure than women and, independently of exposure, also have higher risks of a bacterial illness, then the above model might give a statistically significant positive value of K due to the statistical *association* between risk and exposure, even if changing exposure would have no effect on risk.) Therefore, these methods are not recommended. However, structural equation modeling (SEM) and other techniques of causal analysis (Greenland and Brumback, 2002) can sometimes be used to develop statistical regression equations with valid causal interpretations ("structural" equations) and can produce suitable estimates of the structural (not reduced-form) coefficients. When a simple regression approach can be justified, it can potentially save a great deal of effort

4.5 Rapid Risk Rating Technique (RRRT) Approach

Although population-level regression model are usually not appropriate for risk assessment, they suggest a practical approach to quantitative risk assessment that abandons attempts to assess exposures and dose-response models separately and instead focuses on directly predicting the *risk-per-serving* for food commodities under different risk management interventions. Conceptually, the following linear regression model:

Risk = Background risk + K*Exposure

could provide an adequate basis for assessing the impact of exposure on risk *if* there were a way to estimate K as a causal parameter (so that each unit change in Exposure would cause a K-unit change in Risk) and *if* the relation were truly linear, as in this model. For example, if the Exposure variable were known to cause a fraction F of annual illness cases, then one might invert F = [(K × Exposure)/(Background + (K × Exposure)] to obtain K = [F/(1 – F)] × (Background/Exposure). While more is clearly needed, e.g., to account for multiple sources of exposure and nonlinearities in the exposure-response relation, this basic idea of estimating causal coefficients from fractions of cases caused by specific sources appears worthwhile. To achieve the benefits of such a simple approach, it can be developed as follows:

1. Define "Risk" for an individual as the expected number of illnesses per capita-year from the hazard of interest (e.g., of campylobacteriosis cases from chicken servings) for that individual. Define the relevant "Exposure" variables as the servings-per-year of various types (e.g., from fresh chicken, frozen chicken, etc.).
2. Allow for a background level of exposure that is not affected by the interventions being evaluated.
3. Assume that a risk management intervention, such as a change in animal drug use, can affect several components of total exposure (e.g., servings-per-year of chicken from airsacculitis-positive flocks and servings-per-year of chicken from airsacculitis-negative flocks). Thus, total exposure (e.g., measured in infective doses ingested per year) is decomposed as a sum of exposures from particular sources that can be affected by risk management interventions (e.g., resistant *vs.* susceptible bacteria in servings from ill *vs.* well flocks.)
4. Estimate the expected *change in risk* from a specific intervention that changes one or more of the exposure terms, as follows:

   $\Delta\text{Risk} = K_1(\Delta\text{exposure of type 1}) + K_2(\Delta\text{exposure of type 2}) + \ldots$

   Here, (Δexposure of type j) denotes the change in exposure of type j caused by an intervention. This will usually be a change in the number of servings of a given type ingested per year. Similarly, K_j denotes the average *risk-per-serving* for servings of type j (or, more generally, the risk per unit of exposure of type j. We will use servings as the most important interpretation.) The coefficient K_j is interpreted causally; thus, it should include indirect effects from cross-contamination of other foods due to servings of type j, but it should not reflect non-causal associations between exposure to type j servings and resulting risk. The intended interpretation of K_j is that it reflects the *average change in risk*

per unit change in exposure of type j. (For small changes in exposure, $K_j \approx \partial(\text{Risk})/\partial x_j$, where x_j denotes the amount of exposure of type j. However, practical applications often deal with large changes. For example, if banning an animal antibiotic were to remove *all* resistant bacteria of a certain type in animal food products of type j, without affecting any of the other exposure terms, then the relevant value of K_j*Exposure$_j$ would be the entire reduction in risk that would occur when Exposure$_j$ is set to zero.) For linear models, traditional structural equations modeling (SEM) (Greenland and Brumback, 2002) provides an appropriate statistical approach for estimating the coefficients (and the background term) from exposure-response data. In general, however, nonlinearities must be expected.

5. To deal with the fact that many dose-response relations are not linear (e.g., Figure 1), we interpret the linear coefficients K_j as giving *bounds* on the true changes in risk caused by changes in the corresponding exposures. For example, if the true (but perhaps unknown) dose-response relation is convex (upward-curving) over the exposure range of interest, and if the current level of exposure of type j, denoted x_j, causes a risk of R_j that could be prevented by setting x_j to zero while holding all other exposures fixed, then estimating K_j as the slope of the line from (0, 0) to (x_j, R_j), i.e., estimating it as $K_j = R_j/x_j$, provides a plausible *upper bound on the reduction in risk* that would be achieved by reducing exposure of type j (by assuming that *all* of the current risk level caused by x_j would be removed). But it provides a plausible *lower bound on the increase in risk* that would be caused by an increase in x_j. To draw confident conclusions comparing the changes in risk caused by different interventions, we can exploit such bounds derived from the linear-model approximation (i.e., Risk = Background + $\sum_j K_j x_j$) to the true (perhaps multivariate nonlinear) exposure-response relation.
6. As a starting point for estimating K_j, assume that K_j is proportional to the fraction of total illness-causing exposures contributed by source j. "Total illness-causing exposures" coincides with total exposure (CFUs ingested per year) of the bacterium if the dose-response relation is linear with no threshold. It is the fraction of servings with loads above a certain threshold if the dose-response function is threshold-like. (The background term, which can be denoted K_0, corresponds to an "other" category.) Thus, the initial estimate of K_j becomes:

   (total cases per year) × (fraction caused by exposures from all servings) × (fraction of illness-causing exposures from servings of type j).

For initial estimates, linear no-threshold dose-response and/or threshold-type dose-response relations might be assumed, with different non-linear dose-response relations being tried in sensitivity analyses to see how robust the conclusions are to uncertainties in the dose-response relation.

7. If the coefficients are significantly different for different subpopulations (e.g., with immunocompromised individuals having higher risks than others as in Table 1), then the preceding steps can be applied to each such subgroup separately (data permitting). The total impact of an intervention in the population is then found by summing the estimated impacts for all subgroups.

When the required fractions can be estimated and convexity of the dose-response relation can be assumed, the preceding framework is intended to allow for rapid estimation of the direction and sizes (or bounds on sizes) of changes in human health effects that are likely to be caused by interventions that change exposures. Because the assessment of K_j is based on the relative contributions made by different exposure sources (e.g., serving "types" and background) to total illness-causing exposures, the exposure modeling methods discussed in Chapter 4 (e.g., simulation-based, attribution-based, and systems dynamics-based) can be used in the assessment.

The sum-of-products modeling approach also has a systems dynamics motivation. Suppose that human illness dynamics are described as:

$$dIH/dt = [b + K_{IA}IA + K_{HA}(1 - IA)](1 - IH) - rIH,$$

where IH denotes the "ill human" pool, i.e., the fraction of the population that is ill at any moment; (1 – IH) is the susceptible human pool (not currently ill), r is the recovery rate per ill person per unit time, b is a background infection rate parameter, K_{IA} and K_{HA} are exposure-related infection rate parameters (for servings from the ill animal and healthy animal pools, respectively); and IA and 1 – IA denote the fractions of servings from ill and healthy animals, respectively. Then for rare illnesses (for which $IH/(1 - IH) \approx IH$), the steady-state endemic illness level in the human population is proportional to $b + K_{IA}IA + K_{HA}HA$. This provides further motivation for studying the simple linear structural equation model:

Risk = background + $K_{IA} \times$ (servings from ill animals) + $K_{HA} \times$ (servings from healthy animals).

Example: Estimating Risks Caused and Prevented by an Intervention

Setting: To illustrate the RRRT approach, consider the following hypothetical example. Suppose that 57% of campylobacteriosis cases in the U.S. each year are caused by eating *Campylobacter*-contaminated servings of chicken (FDA-CVM, 2001. As discussed in Chapter 4, data such as Stern and Robach, 2003 suggest that 10% might be more realistic, but we will use 57% for purposes of illustration and then examine the sensitivity of the conclusions to this number.) Suppose also that 1% of all campylobacteriosis cases are resistant to a certain human antibiotic, A, (such as erythromycin, CDC, 2000), independent of the source of exposure (chicken or other). Assume that susceptible cases cause an average of 6 illness-days while resistant ones cause an average of 8 illness-days. The current microbial load frequency distribution on chicken servings, relative to the dose-response relation, is unknown, but, it is known that most chicken servings have either zero or very low CFU counts, while a small fraction carry much higher loads that account for a disproportionate (at least ten-fold) increase in risk. (For example, the high-risk servings might come from airsacculitis-positive or necrotic enteritis-positive flocks, or from flocks with other bacterial illnesses or conditions that lead birds to fail to reach the normal size and weight range; see e.g., Russell, 2003 and Dawe, 2004.)

Problem: A proposed ban on selected current antibiotic uses in chickens would increase the prevalence of chicken servings with relatively high microbial loads from 0% before the ban to at least 1% after it, while reducing the prevalence of A-resistant bacteria in chicken servings by an unknown amount. Assess the probable net human health impact of this intervention, expressed as a fractional change in the illness-days per year from campylobacteriosis in the population. How robust is the conclusion to uncertainties?

Solution: Prior to the intervention, each case causes an average of $(0.01) \times (8) + (0.99) \times (6) = 6.02$ illness-days. To obtain an estimated *upper bound* on the human health *benefits* of the proposed ban, assume that it eliminates *all* A-resistant strains. This reduces the average illness-days per case from 6.02 to 6.00. To obtain a *lower bound* on the incremental human health *risks* from the ban, assume that the dose-response function is linear, as follows:

E(cases per year) = Background + K_1(exposure from low-risk servings) + K_2(exposure from high-risk servings) = Background + K × [number of low-risk servings + 10 × (number of high-risk servings)],

where K is the slope of the linear no-threshold dose-response function plotting expected illnesses against CFUs ingested. (This linear approximation will provide a lower bound on the true but unknown incremental illnesses caused by increased

exposure, if the true dose-response function is convex.) Dividing through by the baseline (pre-intervention) levels gives:

Relative number of cases = 0.43 + 0.57 × (relative exposure),

where the "Relative number of cases" is defined as 1 for the baseline and "relative exposure" is also defined as 1 for the baseline. (This simplification normalizes and non-dimensionalizes the calculation, as is common practice in applied mathematical modeling.) Following the intervention, relative exposure increases from 1 to at least (0.99) × (1) + (0.01) × (10) = 1.09, assuming that the average microbial load in high-risk servings is at least ten times the average microbial load from other servings. Thus, the estimated lower bound on the relative number of cases per year following the ban would be:

Relative number of cases = 0.43 + 0.57 × (1.09) = 1.0513,

corresponding to about a 5% increase in cases per year. Combining this with a 6.00/6.02 = 0.9967 reduction factor in the average illness-days per case yields:

Relative number of illness-days following the ban = 1.0513 × 0.9967 = 1.048.

Thus, the bounding assumptions used here lead to the conclusion that the proposed ban would increase the population risk (illness-days per year) by a factor of *at least* 1.048, corresponding to a 4.8% increase in campylobacteriosis illness-days per year in this population.

If only 10%, rather than 57%, of cases are caused by consumption of chicken servings, then the calculations are revised as follows:

Relative number of cases = 0.90 + 0.10 × (1.09) = 1.009
Relative number of illness-days following the ban = 1.009 × 0.9967 = 1.006.

Thus, the qualitative conclusion that the proposed ban would harm human health is robust to this change in the assumed fraction of cases caused by chicken consumption. It is also robust to uncertainty about the true fraction of resistant cases caused by antibiotic use in chicken (since we used an extreme bounding assumption that the proposed ban would eliminate *all* resistance; hence, if the truth is less extreme, it will only reinforce the conclusion.) Similarly, it is insensitive to uncertainties about the true shape of the dose-response function, provided that the function is convex (upward-curved). However, the finding of an increased risk is sensitive to the problem assumption that the proposed ban will increase exposure to high-risk servings by at least 1%.

4.6 Validation of Risk Characterization Results

The predictive risk assessment models on which risk characterizations are based should be validated before the risk characterizations are accepted as useful for guiding risk management decisions. The exposure assessment and dose-response sub-models, as well as the full predictive risk assessment model formed by composing them, may be validated by comparing their predictions to observations in data sets not used in creating the models. Validation data typically come from multiple distinct populations (e.g., in different geographic sub-regions, seasons or years) or from multiple individual exposure, covariate, and response histories when these data are available (e.g., in case-control studies). Model-predicted risks for each subpopulation or individual are compared to observed illness rates to determine whether the observed values could plausibly have been drawn from the risk distributions predicted by the model. Formal goodness-of-fit tests and model diagnostics (Greenland, 1989) are used to compare observed and predicted values.

Example: Validation Testing of a Simple Risk Model

Suppose that the risk of campylobacteriosis from chicken were hypothesized to be proportional to the quantity of contaminated chicken servings consumed, and that this quantity, in turn, is hypothesized to be a constant fraction of total chicken servings. Then the resulting simple risk assessment model,

$$\text{Risk} = \text{Background} + K^*(\text{servings of chicken consumed}),\ K > 0,$$

could be compared to real data (e.g., from the case-control studies of Effler *et al.*, 2001 or Friedman *et al.*, 2004) to assess the validity of the testable prediction that probability of being a case rather than a control increases with servings of chicken consumed. Perhaps surprisingly, such validation tests suggest that the preceding simple model is *not* valid: instead, as discussed in Chapter 4, risk of campylobacteriosis generally *decreases* significantly with quantity of chicken consumed (Cox, 2002; see also Table 1 of Effler *et al.*, 2001). Thus, the above simple conceptual model should be revised to focus on exposures that are positively associated with campylobacteriosis risks (such as restaurant- and commercially-prepared meats, including chicken in restaurants; or perhaps just those chicken servings that have exceptionally high microbial loads, such as those from airsacculitis-positive chicken flocks.) The conceptual model in which all chicken servings are positively associated with increased risk is refuted by several different case-control data sets, and therefore should not be used for risk assessment.

If the predicted risks do *not* match the validation data, then model inputs, assumptions, and functions should be checked and revised. Often, results of model-checking and diagnostics (Greenland, 1989) can be used to refine the initial model. Differences between predicted and observed values suggest changes in the model that will explain and reduce the differences. If a model uncertainty decision tree for uncertainties (Sielken *et al.*, 1995) has been used to organize and display modeling uncertainties, then the weights of evidence for different branches may be updated to increase the relative weights on branches (i.e., assumption sets) that yield predictions that are most consistent with the validation data.

On the other hand, if the model-based predictions *do* adequately match validation data, as indicated by goodness-of-fit tests and model diagnostics (e.g., plots of residuals), then the underlying risk model may be used to make predictions for risk assessment within the validated range of conditions. Uncertainty and sensitivity analyses should then be used to show how robust the scientific and policy-relevant conclusions of the model are to uncertainties in assumptions (e.g., model structure) and input data.

5. UNCERTAINTY AND SENSITIVITY ANALYSES

5.1 Uncertainty Analysis

Uncertainty analysis is used by risk assessors to characterize both uncertainty and variability in risk estimates. These are distinct concepts (Hoffman and Hammonds, 1994). *Uncertainty* about risk reflects the width of the range of risk estimates that are considered plausible in light of available data, and hence reflects the extent of ignorance about the correct value. For example, Bayesian posterior distributions, conditioned on available data, for the parameters of (a) a frequency distribution, n(x) of microbial loads, x, in servings of a food commodity; and (b) a parametric dose-response model r(x), induce a corresponding posterior distribution for expected Risk = $\Sigma_x r(x)n(x)$. The width of the shortest interval of Risk values that contains 95% of this posterior distribution then constitutes a 95% *Bayesian uncertainty interval* for the true but unknown value of $\Sigma_x r(x)n(x)$. Uncertainty intervals, whether Bayesian, classical (e.g., confidence intervals), or simulation-based (e.g., Monte-Carlo approximations of output distributions based on input joint distributions and model assumptions) can potentially be narrowed, at least in principle, by obtaining better information. By contrast, *variability* in risk reflects true differences in individual exposures or dose-response functions. It refers to the width of the frequency distribution of individual risks in a population. Hence, it cannot be reduced

or eliminated by collecting additional information, as better information will only show more clearly the variability in individual risks within the population. For microbial risks, identifying the sizes and relative susceptibilities (as in Table 1) of different subpopulations helps to characterize variability. If exposures also differ among subpopulations with different susceptibilities, then these differences contribute to the final frequency distribution of individual risks.

Uncertainty about risk estimates is often described by upper and lower 95% statistical confidence limits, derived from corresponding 95% confidence intervals or joint confidence regions for the underlying model parameters and inputs (Haas *et al.*, 1999). For example consider the simple linear structural equation risk model

$$\text{Relative number of cases per year} = B + K \times (\text{Relative exposure}),$$

where the coefficients are normalized so that $B + K = 1$ (with B and K both non-negative) and the variables are normalized so that initially Relative number of cases per year = Relative exposure = 1. If K has a 95% statistical confidence interval ranging from 0.1 to 0.5, then the corresponding 95% confidence interval for the relative number of cases per year following a doubling of the initial exposure level would be:

$$[0.9 + 0.1 \times 2,\ 0.5 + 0.5 \times 2] = [1.1,\ 1.5].$$

If the relative exposure created by an intervention were also uncertain, then the confidence interval for risk would reflect the joint confidence region for K and relative exposure.

Uncertainty about risk can also be described by *uncertainty intervals* around point estimates (or by *uncertainty bands* around entire curves, as in Figure 3). These are typically derived by fitting risk models to many different randomly selected subsets of data or to many different Monte Carlo simulation runs and then keeping the intervals (or bands) that contain 95% of the resulting estimates. The most detailed outputs of quantitative uncertainty analyses are joint posterior probability distributions for model quantities and predictions after conditioning on observed data (and on modeling assumptions). These can be displayed as joint confidence regions for model parameters and/or predictions.

Important computational methods and algorithms for uncertainty analysis include:

- *Monte Carlo uncertainty analysis* using commercial software products such as Analytica™, @RISK™, Crystal Ball™ (Vose, 2000). (For more on uncertainty and sensitivity analysis software, see the

descriptions at product web sites. Without endorsing particular products, we note that the following sites contain useful tutorial and expository information:
www.palisade.com/html/risk/new_in_risk45.html
www.decisioneering.com/cb_features.html
www.merak.com/kr/files/316/DTreeAboutthisrelease.pdf.)

- *Bayesian uncertainty analysis* for estimation of joint confidence regions for model parameters and predictions (e.g., using the free WINBUGS software for Markov Chain Monte Carlo Bayesian inference with missing data.)
- *Bootstrapping and other resampling techniques* for estimating joint confidence regions for model parameters and predictions.
- *Model cross-validation* techniques for estimating the accuracy and prediction error characteristics of model predictions from performance on multiple subsets of data.

These methods are discussed in computational statistics texts and in risk analysis texts such as Haas *et al.*, 1999, Vose, 2000, and Cox, 2001.

Example: Monte Carlo Uncertainty Analysis of a Product of Factors

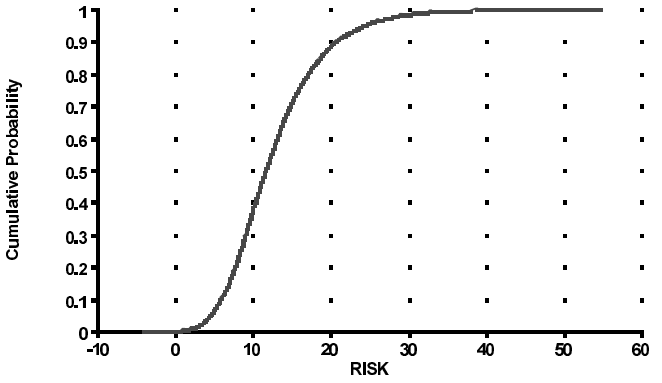
To illustrate one of the key concepts of uncertainty analysis, Figure 7 displays the uncertainty in a risk estimate calculated using the RRRT formula:

$$Risk = Exposure \times Dose\text{-}Response \times Consequence$$

where: Risk = expected number of excess illness-days per year, Exposure is measured in potentially infectious meals ingested per year in a population, Dose-Response = expected number of illnesses caused per potentially infectious meal ingested, and Consequence is measured in illness-days caused per illness.

For purposes of illustration, the point estimates are taken to have median values of: Exposure = 10, Dose-Response = 0.2, and Consequence = 6 days. To express uncertainty, Exposure is modeled as a log-normal distributions with a geometric standard deviation of 1.4; Dose-Response is modeled as a Bernoulli random variable having value of 1 with probability 0.2 (for "susceptible" members of the exposed population) and a value of 0 otherwise (and hence a mean value of 0.2); and Consequence is modeled as a normally-distributed random variable with mean of 6 days and standard deviation of 2 days. The curve is a cumulative probability distribution for Risk, given these uncertain estimates of Exposure, Dose-Response, and Consequence. It was generated using the Analytica™ Monte Carlo uncertainty analysis software (http://www.lumina.com/). Rai and Krewski, 1998, discuss other methods of uncertainty and variability analysis for such multiplicative risk models.

Figure 7: Cumulative Probability Distribution for a Product of Factors



The increasing availability of high-quality commercial software tools for uncertainty analysis of risk models has made it possible for risk analysts to rely on these tools for sophisticated resampling, conditioning, and Monte Carlo simulation algorithms and uncertainty analysis displays. Thus, without needing to know the details of the algorithms used to generate uncertainty bands, distributions, and intervals, most risk assessment practitioners can now easily create not only "1D" Monte Carlo uncertainty analyses (which conflate uncertainty and variability analyses by sampling individual attribute values and then simulating individual risks conditioned on these values), but also "2D" or "second-order" Monte Carlo uncertainty analyses. These separate uncertainty analysis and variability analysis by quantifying uncertainty bands and intervals for risk *conditioned* on individual attributes, while separately quantifying the joint frequency distributions of individual attributes… and uncertainty about this joint frequency distribution, i.e., about variability. Figure 2 of Chapter 2 illustrates a typical "1D" presentation of uncertainty about risk, in the form of a simulation-based distribution for the number of illnesses per year predicted by a risk model. Figure 3 of Chapter 2 shows uncertainty about risk conditioned on alternative modeling assumptions. Table 4 of this chapter indicates where 2D and 1D analyses were used in the *Listeria* risk assessment. For conceptual clarity in distinguishing between uncertainty and variability, it is often useful to interpret a risk assessment as consisting of separate component risk assessments (including uncertainty analyses) for each of several distinct subpopulations, thus capturing the variability of risks across homogeneous subsets of the heterogeneous population. Total risk can be viewed as the superposition of these component population risks.

5.2 Sensitivity Analysis

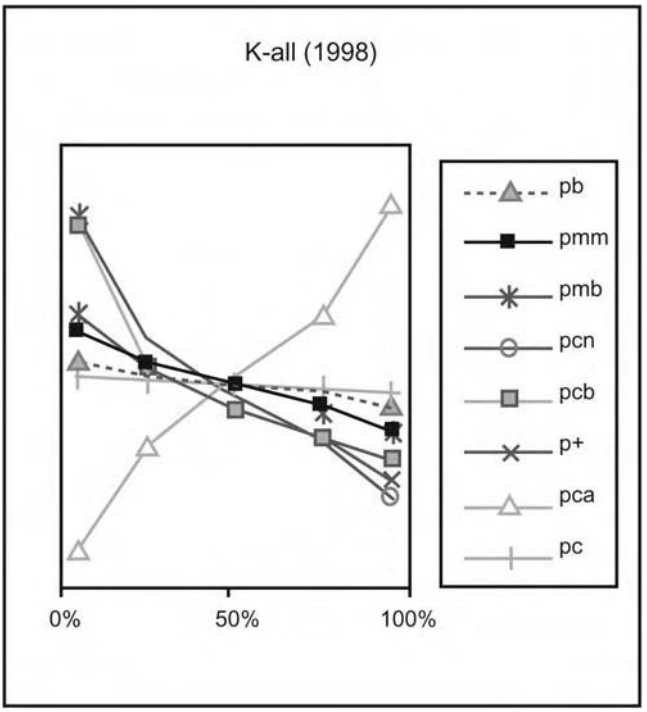
Sensitivity analysis (Saltelli *et al.*, 2004) identifies the inputs to a risk assessment that most affect its outputs. It also shows how sensitive the outputs are to plausible variations in inputs, and how risk estimates (both point estimates and uncertainty intervals) and recommended risk management decisions change as the inputs are varied and as current uncertainties are resolved in different ways. Several types of diagrams have been developed for these purposes, as illustrated in the examples in this section. Most commercial software tools that create Monte Carlo uncertainty analyses for risk assessment also generate sensitivity analysis displays. For example, modern decision analysis software tools (Clemen, 2000) typically include sensitivity analysis diagrams and Value-of-Information (VoI) analysis capabilities to show how and whether more information is likely to change current decision recommendations.

Example: Parameter Sensitivity Analysis Using Spider Diagrams

In the late 1990s, the US FDA's Center for Veterinary Medicine hypothesized that the risk of domestically-acquired fluoroquinolone-resistant campylobacteriosis in the US population (not including cases linked to human use of fluoroquinolones) increases in direct proportion to pounds per year of *Campylobacter*-contaminated chicken consumed (FDA, 2001). Figure 8 is a *spider diagram* sensitivity analysis plot from that analysis, showing how the proportionality constant, K_{all}, varies with different model input parameters. (The horizontal axis shows deviations of individual parameters around their baseline levels. The vertical axis shows resulting deviations of K_{all} around its baseline level. The vertical axis need not be labeled because all values are expressed relative to the baseline value.) The parameters of the model included:

- Proportion of *Campylobacter* infections from chicken that are fluoroquinolone resistant (p_{rh});
- Probability a case of campylobacteriosis is attributable to chicken (p_{ca});
- Probability that a stool will be requested and submitted from a patient with non-bloody diarrhea (p_{cn}); and
- Probability that the culture will confirm *Campylobacter* given it was tested ($p+$)
- Other parameters dealing with the probability of seeking medical treatment, being prescribed a resisted antibiotic (a fluoroquinolone), etc.

Figure 8: A "Spider Diagram" Deterministic Sensitivity Analysis Plot



Source: http://www.fda.gov/cvm/antimicrobial/RRAsec5.pdf

This plot was explained as follows (FDA, 2001):

"[The figure] illustrates the parameters that contribute the most to the ratio K_{all}. The parameters p_{ca} and p_{cn} produce the greatest vertical range and therefore are the most influential input parameters. … The parameter p_{ca} is the only significant parameter plotted that contributes to the uncertainty from modeling contamination of chicken meat, i.e. all the other parameters correspond to determining the human health impact which means that we have more uncertainty about the human health side than the broiler side." (http://www.fda.gov/cvm/antimicrobial/RRAsec5.pdf)

Figure 8 illustrates some of the strengths and limitations of sensitivity analyses based on varying model input parameters and plotting how the output changes. It is useful in indicating the sensitivity of the hypothesized risk relation to p_{ca}, the judged "Probability a case of campylobacteriosis is attributable to chicken". However, the plot does not reveal sensitivities to model uncertainties and errors. For example, if the fundamental hypothesis that "Risk = K_{all} × Exposure" is incorrect (e.g., because low levels of exposure reduce risk while only very high

levels increase it), then no sensitivity analysis of K_{all} will illuminate the errors to be expected in predictions from the model. (Such model specification errors and model uncertainties can be addressed by other methods, such as nonparametric models, model cross-validation error estimates, and Bayesian model averaging, surveyed in Cox, 2001, Chapter 3). Similarly, sensitivity analyses do not necessarily reveal use of erroneous models and formulae. For example, suppose that the quantity "Probability that a resistant case of campylobacteriosis is attributable to chicken" was estimated as the product of "Probability a case of campylobacteriosis is attributable to chicken" × "Probability that a case is resistant". This formula is incorrect. (For example, it implies that if *all* campylobacteriosis cases are caused by chicken and 10% of them are resistant, then 100% = 100% × 10%.) Yet, sensitivity plots can be produced without noticing such formula errors. Thus, they may provide an unwarranted sense that relevant uncertainties have been accounted for, while not revealing mistakes in the underlying assumptions. (This risk assessment was used by FDA to recommend that fluoroquinolones be withdrawn from use in poultry in the United States. FDA considered its results convincing, in part because of the elegant supporting sensitivity analyses, even though some critics noted that they were based on erroneous formulas and therefore lacked validity.)

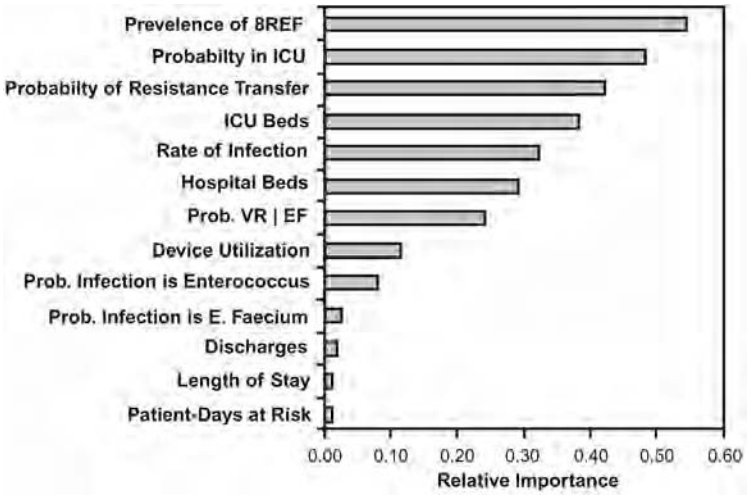
Example: Relative Importance Plots

Figure 9 illustrates a "relative importance plot" from an FDA-CVM draft risk assessment of the animal antibiotic virginiamycin (VM) (FDA-CVM, 2004). VM has been used to promote growth and control diseases in food animals (such as necrotic enteritis in chickens), but it is also almost identical to the human antibiotic Synercid™ (quinupristin-dalfopristin ("QD"), a combination streptogramin drug sometimes used to treat vancomycin-resistant *E. faecium* bacteremias among intensive care patients (see Chapters 6-8).

The FDA-CVM draft report explains the sensitivity analysis plot in Figure 9 as follows:

> "One of the benefits of the risk assessment process is that key data gaps are usually identified suggesting areas of needed research and generating testable hypotheses. Toward this end, a simplified sensitivity analysis was performed… using built-in features of Analytica® [the commercial risk analysis tool used to perform the quantitative risk assessment]. A "variable importance" is the absolute rank-order correlation between the sample of output values and the sample for each uncertain input. Importance analysis 'is a robust measure of the uncertain contribution because it is insensitive to extreme values and skewed distributions.

Figure 9: Relative Importance of Variables in Virginiamycin Risk Model



(*Source:* http://www.fda.gov/cvm/antimicrobial/antimicrobial.html)

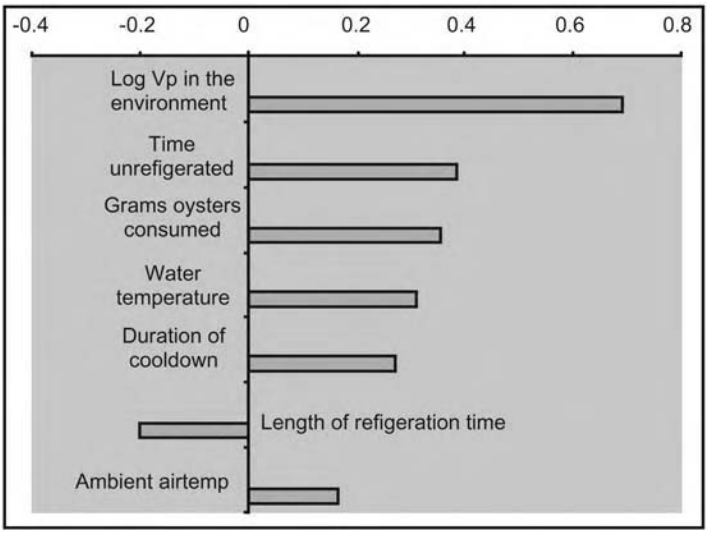
> Unlike commonly used deterministic measures of sensitivity [such as those in the preceding example for fluoroquinolones], it averages over the entire joint probability distribution. Therefore, it works well even for models where the sensitivity to one input depends strongly on the value of another.' (Analytica® documentation)."

Example: A Tornado Diagram for Factors Affecting Risk Per Serving

Figure 10 shows a "tornado diagram" indicating the fractional change in predicted risk-per-serving of the summer harvest of Louisiana Gulf Stream raw oysters as each of several input factors, listed on the left side of the diagram, varies over a range of plausible values. When all input factors have the levels assumed in the base case (i.e., if all of these point estimates are correct), then the model-estimated risk also has its base case value, and so the fractional change from this nominal value (indicated by the horizontal scale at the top of the diagram) is zero. If the "time unrefrigerated" increases to the largest plausible value considered (e.g., to the highest value in a 95% confidence interval of uncertainty interval for this input parameter), then the model-predicted risk would increase by almost 40%. Conversely, increasing "length of refrigeration time" could reduce risk-per-serving (relative to the base case) by more than 20%. The other horizontal bars are interpreted similarly.

While tornado diagrams are widely used in decision and risk analysis to give an impression of the sensitivity of model outputs to model inputs (see e.g., Clemen and Reilly, 2000), they have several limitations. One is that, as in Figure 10, it is not

Figure 10: Tornado Diagram of Influential Parameters Affecting Predicted $\log_{10}$ Risk of *V. Parahaemolyticus* Illness per Serving of Raw Oysters



Source: FDA-CFSAN 2001 http://vm.cfsan.fda.gov/~dms/vprisk6.html

always clear upon inspection exactly what ranges were considered for each input. In addition, the diagrams vary one input at a time, and thus do not illuminate sensitivities to interactions among inputs. (Classification tree analysis of predicted output values *vs.* combinations of input values provides one method for studying sensitivities to such interactions.) Also, the main drivers of uncertainty may not be revealed in the diagram. For example, as the discussion in FDA-CFSAN 2001 explains, the chief source of uncertainty in this example is uncertainty about the correct dose-response function. This is not revealed by the bars in Figure 10. More generally, sensitivity analyses usually do not address what has been termed "modeler uncertainty", meaning "difference in problem formulation, model implementation, and parameter selection originating from subjective interpretation of the problem at hand" (Linkov and Burmistrov, 2003).

On the other hand, tornado diagrams can be very helpful in informing risk management decisions. One use is to partition the horizontal bars of the diagram into contiguous intervals, each a different color, indicating which risk management alternative (e.g., to introduce a new animal antibiotic use *vs.* not; or to terminate an existing use *vs.* allowing it to continue) is predicted to be "optimal" (e.g., maximizing expected utility, given the risks predicted by the risk assessment model) as each input is varied. This allows one to see at a glance by how much each input would have to deviate from its assumed baseline level to change the recommended decision. For further development of the decision-analytic uses of sensitivity analyses, see Clemen and Reilly, 2000.

In summary, sensitivity analyses show how risk estimates, uncertainty intervals, and recommended decisions vary as inputs are changed and as uncertainties for inputs are resolved in different ways. Sensitivity analysis should also show how model results and predictions change if different plausible modeling assumptions are made. It is important to include the impacts of such variations, as well as of data uncertainties and parameter estimates, in any complete risk assessment. Technical methods for displaying the results of sensitivity analyses (e.g., tornado diagrams, relative importance plots, spider diagrams) have been developed in the decision and risk analysis literatures and are appropriate to include in technical discussions and presentations of risk analysis results. These displays help professional risk analysts and quantitatively oriented decision-makers to understand and communicate which inputs to a risk model are most uncertain and contribute most to uncertainties in the model's predictions.

However, characterizing uncertainty about model assumptions (e.g., about the mathematical form of a model's equations and which inputs affect the model's outputs and other variables) requires more powerful statistical methods such as those mentioned above (Cox, 2001, Chapter 3.) Meanwhile, *if* a model is assumed to describe the true causal relations among its variables with useful accuracy, *then* relatively simple plots such as those above suffice to identify the key drivers of uncertainties in its results.

Much more can be done with analytic methods of sensitivity analysis for certain parametric families of models (Saltelli *et al.*, 2004). As a mathematically trivial but practically important special case, if risk is expressed as a product of factors, then over-estimating or under-estimating any one factor creates a proportional over- or under-estimate of risk, respectively, with the constant of proportionality being the product of the remaining factors. More generally, if the effect of exposure on risk can be expressed as the difference of two products of powers of exposure variables, as in S-system theory for biological systems theory, then mathematical expressions and algorithms are available for analyzing the sensitivity of system dynamic responses and steady-state equilibria to changes in inputs (e.g., Cascante *et al.*, 1989).

Given the increasing availability of high-quality uncertainty and sensitivity analysis software, risk analyses should now be expected to quantify and present all key sensitivities, show estimated variability of risk metrics in the exposed population, and provide uncertainty analysis displays for their major conclusions. In particular, *model uncertainties* (e.g., about the shape of the microbial load exposure distribution, the dose-response relation, the selection and coding of variables in multivariate models, and the effects on risk of individual covariates and their interactions) should be

discussed and sensitivities to these aspects of modeling should be displayed. For example, instead of selecting individual models and then reporting parameter estimates and uncertainties assuming that the selected model is correct, techniques such as Bayesian model averaging (BMA) can be used to better characterize the plausible set of models and risk estimates that are most consistent with the data (Viallefont *et al.*, 2001).

6. DISCUSSION AND SUMMARY

This chapter has described dose-response modeling and risk characterization for microbial and antimicrobial risk assessment. The best established approach to quantitative risk assessment, following the National Academy of Sciences (NAS) framework (Jaykus, 1996), develops separate estimates for n(x), the number of servings-per-year of the food commodity ingested with dose x of a bacterial pathogen; and r(x), the conditional probability of illness given ingestion of a serving with dose x. *Exposure assessment* modeling (Chapter 4) provides the exposure frequency distribution n(x) for different x (or for x values binned by order of magnitude, as in Table 3.) *Dose-response modeling* provides the dose-response model r(x) for each x. *Risk characterization* combines these components to estimate the frequency distribution of the number of illnesses per year: the mean risk = expected number of cases per year = $\sum_{\text{all } x} n(x)r(x)$. The actual number of cases per year is the sum over all x values of binomial random variables with parameters n(x) and r(x), denoted Bin[n(x), r(x)]. That is, the random number of cases-per-year is: $\sum_{\text{all } x} \text{Bin}[n(x), r(x)]$, for any homogeneous population having the same r(x) for all individuals.

Both the exposure frequency distribution n(x) and the dose-response function r(x) may differ for different subpopulations, reflecting *variability* in these two components and implying variability in the resulting risk estimated from them. If they are expressed as parametric functions [such as $n(x) = M*\lambda e^{-\lambda x}$, where M is the total number of servings and λ is a parameter to be estimated; and r(x) is the Beta-Poisson dose-response model with parameters α and β], then *uncertainty* about the parameter values induces uncertainty about risk. Monte Carlo uncertainty analysis software is used to calculate and display uncertainty and variability information; "second-order" Monte Carlo uncertainty analysis also quantifies uncertainty about variability. Risk management interventions typically change the exposure frequency distribution n(x), and hence change the probability distribution of risk.

An alternative approach to risk assessment is to express risk as a sum of products of factors:

$$\text{Risk} = \text{Background} + (K_1 \times \text{Servings}_1) + \ldots + (K_S \times \text{Servings}_S),$$

where S is the number of distinguished sources of exposures (here assumed to be different types of servings of food commodities) potentially affected by proposed risk management interventions. In this model, Risk refers to the expected number of cases per year. Servings_j is the number of servings per year of type j, such as servings of fresh chicken from airsacculitis-positive *vs.* airsacculitis-negative flocks (or even just "high-risk" *vs.* "low-risk" servings, if there is a bimodal distribution of risk-per-serving). In this framework, the linear equation is interpreted as an approximate structural equation model (Greenland and Brumback, 2002). It provides *lower bounds* for the effects of increasing exposures of each type and upper-bound estimates for the effects of eliminating them (since it implies that *all* of the corresponding risk would disappear) whenever the true (but perhaps unknown) dose-response function is convex. Thus, if this simple linear model predicts that an intervention that increases some exposures and eliminates others has the net effect of increasing risk, then this conclusion should only be strengthened by using a true nonlinear convex dose-response function in place of the linear approximation. For example, if a ban on a current animal antibiotic would increase exposure to susceptible bacteria and decrease exposure to resistant bacteria and if the linear model shows that

> $K_1 \times$ (increase in exposure to susceptible bacteria) $\geq$ $K_2 \times$ (current exposure to resistant bacteria).

then using the true (but perhaps unknown) convex dose-response relation should preserve the inequality:

$$\begin{aligned}\text{Increase in Risk} \geq\ & K_1 \times (\text{increase in exposure to susceptible bacteria}) \\ & - K_2 \times (\text{current exposure to resistant bacteria}).\end{aligned}$$

We therefore refer to this linear modeling method as the *bounding structural equation model* (BSEM) approach.

The coefficients K_j are intended to have direct causal interpretations as the ratios of increases in Risk per unit increase in servings of each type. Each coefficient thus combines exposure information (e.g., CFUs per serving) and dose-response information (e.g., expected illnesses per CFU) into a single quantity. Because the coefficients are interpreted causally, they cannot necessarily be estimated by regressing Risk against the numbers of

servings of different types, since regression coefficients are generally not causal (Freedman, 2004). However, if F_j denotes the fraction of all illness-causing exposures (CFUs) contributed by servings of type j (e.g., as estimated from epidemiological data and/or genotyping, serotyping, resistance typing, or other typing data), then inverting the identity

$$F_j = (K_j \times \text{Servings}_j)/[\text{Background} + K_1 \times \text{Servings}_1 + \ldots + K_S \times \text{Servings}_S]$$

yields the following estimate for K_j:

$$K_j = (F_j \times \text{Risk}_0)/\text{Servings}_{j0},$$

where Risk_0 is the baseline risk level (number of cases per year) and Servings_{j0} is the baseline number of servings of type j. Both of these quantities can typically be estimated from available data, and sensitivity analysis can then be used to study how the risk assessment conclusions change as the input F_j is varied. When the BSEM model holds and the baseline risk Risk_0, relative exposure fractions F_j, and changes in proportions of servings of different types caused by proposed interventions can be estimated, it provides a practical basis for rapid estimation of the probable direction of the net change in human health impacts caused by the interventions and for putting a quantitative lower bound on the size of any net increase in expected cases per year (or an upper bound on the size of any net decrease in expected cases per year.)

Uncertainty and sensitivity analyses indicate the potential for more information to change conclusions about risks (including their uncertainties) and comparisons of alternative risk management interventions. Showing the potential for change if more information is obtained can provide affected stakeholders with an incentive to collect additional relevant information if they wish to change current risk estimates and risk management decisions.

Both the traditional risk characterization approach that separately estimates exposure and dose-response components and then integrates them via Monte Carlo simulation (the Risk = $\sum_{\text{all } x}\text{Bin}[n(x), r(x)]$ approach) and the Rapid Risk Rating Technique (RRRT) based on the bounding structural equation model approach can be extended to describe time-varying risks as exposure patterns shift in response to interventions or because of the systems dynamics governing changes in ill animals or other drivers of exposure, as discussed in Chapter 4. Both types of models can be used to describe illnesses from antibiotic-resistant as well as antibiotic-susceptible bacteria. However, in our experience, the RRRT approach is often easier to apply with available data for animal antibiotic risk assessments than the traditional approach. It also avoids model uncertainties about separate exposure

distributions and dose-response functions, n(x) and r(x), that can make the predictions from the traditional approach based on separate assessment of these components uncertain and difficult to validate.

The following chapters further develop and apply the bounding structural equation modeling approach and the Rapid Risk Rating Technique (RRRT) based on it, emphasizing the practical estimation of the coefficients from data, uncertainty analysis when systems dynamics are included, and comparison of the probable net human health impacts of alternative risk management decisions.

This page intentionally blank

Chapter 6

Human Health Risks from Virginiamycin: A Case Study

1. INTRODUCTION

This chapter shows how to estimate a plausible upper bound on the harm to human health that could be prevented by discontinuing a specific current use of an animal antibiotic. This may be interpreted as estimating an upper bound on either the human health *benefit* (i.e., decreased human health risk) from a ban on continued use; or on the preventable *risk* allowed if such a ban is not implemented.

In terms of the simple conceptual risk model

$$\text{Change in risk} = K_1 \times (\text{change in exposure to susceptible bacteria}) + K_2 \times (\text{change in exposure to resistant bacteria}),$$

this chapter focuses entirely on the decrease in human health risks from eliminating *resistant* bacteria, and hence on estimating the corresponding causal parameter K_2 and the exposure to resistant bacteria (and how it would change if a ban were implemented.) Chapter 8 continues the example by considering effects on susceptible bacteria.

To illustrate practical aspects of the estimation and bounding procedures with realistic data and despite realistic knowledge gaps, we take as a case study the human health risks from use of a specific animal drug, virginiamycin (VM). VM is an animal drug of considerable practical interest for which improved risk assessment techniques may be useful in current policy-making decisions. It was banned as a growth promoter in Europe in the late 1990s on what were considered precautionary grounds,

but other countries are still weighing whether the potential benefits of continued use outweight the potential harm (e.g., FDA-CVM, 2004).

The calculations in this chapter illustrate the attribution-based and product-of-factors approaches to estimating exposures discussed in Chapter 4 as alternatives to simulation-based approaches. They are used to develop a plausible upper-bound risk estimate for the human health impacts of VM use in chickens and for the potential human health benefits of banning such use. The practical goal is to estimate a plausible upper bound on the human health risk posed by the use of VM in chickens, in both the United States and in Australia. In Australia, concern about the theoretical possibility of spread of QD-resistance from VM-exposed chickens to humans has led to proposals to reduce use of VM as a growth promoter, despite evidence that no such cases have yet occurred. In the United States, a much higher prevalence of VREFs also suggests a need for caution. The quantitative modeling and analysis presented here is intended to provide the information needed to bound potential risks and benefits from banning VM use in chickens; to approximately quantify the maximum plausible size of the possible human health risks involved; and thus to provide risk assessment and uncertainty information that can help in formulating and choosing among alternative strategies for risk management.

1.1 Background on QD, VRE, VREF, and $VREF_A$

E. faecium are commensal bacteria commonly found in the intestines of humans and of food animals such as chickens, pigs, and cattle. Although they normally pose no health risks when competent immune systems protect their hosts against infections by *E. faecium* and other intestinal bacterial, in severely ill human patients with compromised immune systems, such as leukemia, chemotherapy, transplant, and AIDS patients, these normally harmless bacteria can become life-threatening opportunistic infections unless they are controlled successfully with antibiotics. Vancomycin is the antibiotic most frequently prescribed to treat *E. faecium* infections, but may be ineffective against *E. faecium* that express vancomycin resistance genes. Other antibiotics such as linezolid, daptomycin (Sader *et al.*, 2004), and the streptogramin combination quinupristin-dalfopristin (QD), which are usually highly effective against vancomycin-resistant *E. faecium* (VREF), may then become important treatment options (Critchley *et al.*, 2003). Less effective bacteriostatic agents (e.g., chloramphenicol) are also available, and new antibiotics for treatment of vancomycin-resistant cases (e.g., oritavancin, a glycopeptide, and tigilcycline) are in trial (Linden, 2002).

QD was approved for use in the United States in late 1999. A nearly identical QD compound, virginiamycin (VM), has been used for decades as a growth promoter and to prevent and control bacterial illnesses in farm

animals in the United States and in other countries. Poultry in the United States and elsewhere frequently test positive for QD-resistant *E. faecium* (e.g., Hershberger *et al.*, 2005) raising the theoretical possibility that use of VM in chickens may compromise QD effectiveness in treating human VREF infections if such use promotes the spread of QD-resistant strains from chickens to humans (FDA-CVM, 2004). To date, this hypothesis has proved difficult to test and the potential risk has been hard to quantify, in part because of uncertainty about transfer of resistance genes from chickens to humans in realistic settings. At present, such transfer is an unquantified theoretical possibility – one that may appropriately raise concerns but that provides no quantitative data to help define what would constitute prudent risk management. The high prevalence of QD-resistant *E. faecium* in chickens and its low prevalence in humans suggests that transfer from chickens to humans may currently have little or no detectable impact on human health, but likely future impacts have remained uncertain (McDonald *et al.*, 2001).

A current clinical perspective on VREF infections and resistance is as follows:

> "In the United States and Europe, the VanA-resistance phenotype is reported as the most common phenotype. VanA enterococcal isolates exhibit high-level resistance to both vancomycin and teicoplanin, while VanB isolates have variable resistance to vancomycin and remain susceptible to teicoplanin. …Enterococcal infections often occur in debilitated patients and as part of polymicrobial infections. …The streptogramin combination antibiotic, quinupristin/dalfopristin, is available intravenously for the treatment of *E faecium* infections, but it is not effective against *E faecalis* strains. Linezolid, an oxazolidinone antibiotic, is available orally and intravenously, and it is used to treat infections caused by *E faecium* and *E faecalis* strains. …Once VRE [vancomycin-resistant enterococci, including *E faecium* and *E faecalis* strains] is identified in a medical facility, all clinical enterococcal isolates should be tested for vancomycin resistance." (Donskey and Salata, 2003).

Thus, either QD (formulated as Synercid™ for human medicine) or alternatives such as daptomycin or linezolid can be used to treat the vanA VREF (abbreviated $VREF_A$) infections of concern.

If VM used in chickens increases QD-resistant VREF contamination in food products, thus increasing QD-resistant VREF infections in immunocompromised ICU patients (perhaps following inadequate cooking or handling of hospital food), then more of these patients might have to be treated with alternatives to QD. Since linezolid is usually less harsh and at least as effective as QD, this is not necessarily undesirable. However, for patients who do not respond favorably to linezolid – approximately 7.4% of

VRE patients in a study by Linden et al, 1997 – or to other treatment options such as daptomycin, QD may become the treatment of last resort. QD resistance might then increase the probability of QD treatment failure. Therefore, to the extent that QD use in chickens increases QD-resistant VREF infections in ICU patients, it might also increase the number of cases per year not treated effectively by any currently available antibiotics, leading to excess mortalities or illness-days.

Quantitative risk assessment is needed to determine how large this number of excess treatment failure cases per year is. In the absence of such quantitative risk assessments, opponents of animal antibiotic use in many countries have urged that VM use in chickens (the food animal of main concern) and other food animals be restricted or banned to protect against the perceived but unquantified hypothesized risk to human health (JETACAR, 1999; Wegener *et al.*, 1999; APUA, 2002; WHO, 2003). VM has already been banned as a growth promoter in the European Union

2. RISK ASSESSMENT DATA AND METHODS

To quantify the number of QD treatment failures that could be caused by VM use (or prevented by a VM ban) in chickens over the next five years, one may pursue either of two very different approaches. As discussed in Chapters 4 and 5, *simulation-based strategies* seek to track QD-resistant VREF microbial loads from multiple sources, including chickens, through to people and estimate the number of patients infected and treatment failures that will occur with and without a ban. This approach is consistent with much current farm-to-fork process simulation modeling for microbial and antimicrobial risk assessment. However, it requires many data gaps to be filled with more or less *ad hoc* assumptions, and hence cannot easily give a reliable risk assessment for VM (FDA-CVM, 2004). By contrast, an *attribution-based strategy* starts with the number of QD-resistant VREF cases that occur per year and seeks to quantify the maximum fraction of these cases that could have been caused by VM use in chickens. VREF cases, trends, and genotypes in hospitals in multiple countries have been closely studied and monitored, so this attribution-based approach can more easily be constrained and guided by available data (*ibid*).

For purposes of conservative (i.e., upper-bound) risk assessment, a potential *VM-attributable treatment failure* may be defined to occur whenever a patient (1) has a VREF infection that (2) is resistant to QD and (3) could have come from chicken (e.g., is of a type found in chicken); and (4) the patient is prescribed QD (rather than alternative therapies), but (5) QD therapy fails due to the QD-resistance of the strain (rather than other

causes of failure). The conjunction of these five conditions is necessary for a treatment failure to have been caused by resistance due to use of VM in chickens, even though it is not sufficient (e.g., the strain may be of non-chicken origin). At present, transfer of QD resistance from VM-treated chickens to VREF-infected humans has not been established as a scientific fact (FDA-CVM, 2004). If it happens at all, it is a rare event (McDonald *et al.*, 2001). Accordingly, we adopt a contingent risk assessment approach based on a pessimistic hypothesis: we *assume* that transfer of resistance does occur, and then, contingent on this hypothesis, quantify the plausible maximum frequencies of resulting adverse health outcomes. This is accomplished by estimating the frequencies of simultaneous occurrence of conditions (1)-(5) from available medical data and from recent studies of genogroup frequencies among VREFs identified in animal sources and hospital patients (Willems *et al.*, 2000).

Table 1 illustrates the logic of this conjunction-of-conditions approach. It summarizes the attribution-based, product-of-factors risk assessment model for Australia and shows the mean values estimated for its parameters, the formulas used to approximately quantify uncertainties, and key data sources. These elements are explained and discussed below.

Table 1: Summary of Australia Model Parameter Values for Attribution-Based Risk Estimate

Parameter	Mean Value	Formula	Main Sources
Total VRE cases/quarter in Australian hospitals	4.0	Data-based simulation	JETACAR, 1999; Turnridge 2001
Reduction fractions			
1. VRE that are vanA *E. faecium* (VREFs)	0.22	Beta(18, 65)	Turnridge 2001
2. Proportion of VREFs that are exogenous (i.e., not of known nosocomial origin)	0.17	Uniform(.089, .25)	Bischoff 1999; Austen 1999; Thal 1998
3. Proportion of exogenous vanA VREF that come from chicken	0 to 0.12	0 to Beta(11,78)	Willems 2000
4. QD resistance proportion	0 to 0.009	Beta(1,109)	Eliopoulos 1998
5. QD effectiveness attribution	0.71	N(0.705, 0.0362) + Beta(1,109)	Moellering 1999 Linden 2002
Product of fixed reduction factors:	**0 to 0.000029**	Product of above	
Reductions with Time:		(t = quarters)	
6. QD prescription rate	Reduce 15% semi-annually	0.922^{t}	U.S. market data (AMR, 2001)
7. Reduction in VM resistance in chicken-borne VREFs after ban	Decreases to .32 in 5 years	$e^{(-.0570\ t)}$	DANMAP data 1997-2000

The main idea is to start with *all* vancomycin-resistant enterococci (VRE) cases per quarter and then to estimate the size of the subset that might involve treatment failure due to VM use in chickens, based on the conjunction of conditions 1-5 above. As discussed next, these are cases of QD-resistant vanA VREFs of a type (genogroup) that could have come from chickens and that affect patients who would otherwise have responded favorably to QD treatment. The main work of the risk assessment then consists of identifying data-based estimates of the factors in Table 1. These factors, their estimated values, and supporting data and modeling of uncertainty and variability are discussed in the following sections. Risk estimates are then prepared by recognizing that the frequency of a conjunction of causally related conditions can be expressed as a product of their conditional frequencies (see the "Mean Value" column of Table 1 for point estimates of these factors). In terms of probabilities, without loss of generality, Pr(all of conditions 1-5 hold) = Pr(condition 1 holds) × Pr(condition 2 | condition 1) × … × Pr(condition 5 | conditions 1-4).

2.1 Estimating Total Number of VREF Cases per Quarter

Estimating VREF Cases per Quarter in Australia

Figure 1 plots rates of vancomycin-resistant *Enterococci* (VRE) in Australia from Q3 1994 (the first known case) to Q4 1999 (the interval for which data were available), summed over all institutions, by quarter (JETACAR 1999; Turnidge 2001). Figure 2 shows numbers of cases by quarter and institution. Cases within institutions tend to cluster significantly across adjacent quarters, i.e., cases in one quarter make cases in subsequent quarters more likely, suggesting a contagion process (Cliff and Ord, 1981). VRE has high nosocomial (in-hospital) spread potential and is tracked within the United States as a nosocomial phenomenon (NNIS, 2001). Although it is plausible that VRE can spread between as well as within hospitals (Thal et. al, 1998), a PCR analysis of 69 isolates occurring up to August of 1998 concluded that there was "no direct evidence of interhospital transfer of strains" (Bell *et al.*, 1998). For modeling purposes, therefore, the dynamics of contagion within institutions are here described by a mixture of Markov processes, with correlations among hospitals not being modeled.

The heterogeneity of infection processes at different institutions in Figure 2 can be represented by partitioning them into three clusters, as follows: one for institutions that had no cases in consecutive quarters; a second for all other institutions except "D"; and a third cluster consisting of D alone, which has exceptionally high contagion.

Figure 1: History of VRE in Australia

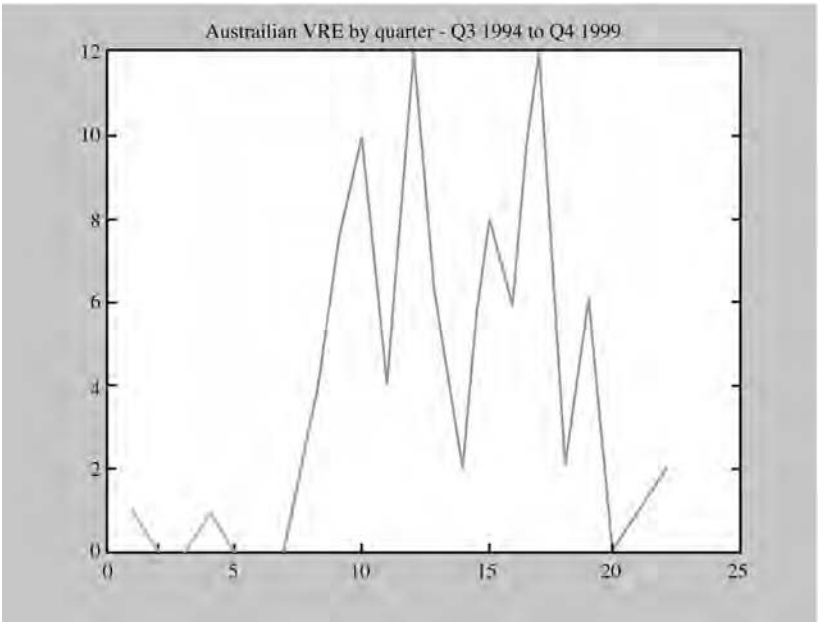


Figure 2: Historical VRE by Institution

	1994		1995				1996				1997				1998				1999				
Institution	q3	q4	q1	q2	q3	q4	q1	q2	q3	q4	q1	q2	q3	q4	q1	q2	q3	q4	q1	q2	q3	q4	**Total**
A	1								1			3											5
B				1																			1
C								1				2	1										4
D								1	1	4	3	2	1			1					1	2	16
E								1		1		1											3
F									1	1									1				3
G									1	1					2								4
H									1														1
I									1	1	1	1				1	1						6
J									1							1							2
K										1													1
L										1		1	2	1	2								7
M												1											1
N												1											1
O													2										2
P														1									1
Q															1								1
R															1								1
S															2		1		3				6
T																1							1
U																1	2						3
V																1	1						2
W																	2						2
X																	2						2
Y																	1						1
Z																	1						1
AA																	1		1				2
AB																			1				1
AC																			2				2
Total	1	0	0	1	0	0	0	3	7	10	4	12	6	2	8	6	12	2	6	0	1	2	83

For each cluster, a Markov transition matrix for transitions between Contaminated, and Not-Contaminated states, can be estimated from the count data in Figure 2. Table 2 summarizes the results.

Table 2: Estimated State Transition Matrices

Cluster 1	**Not Contaminated**	**Contaminated**
Not Contaminated	0.95	0.045
Contaminated	1.00	0

Cluster 2	**Not Contaminated**	**Contaminated**
Not Contaminated	0.91	0.09
Contaminated	0.50	0.50

Cluster 3	**Not Contaminated**	**Contaminated**
Not Contaminated	0.67	0.33
Contaminated	0.25	0.75

Table 2 shows the maximum likelihood estimates (MLEs) for the transition parameters, assuming that observations start at each institution in the quarter in which its first case occurs. (This is a conservative assumption, since it ignores periods without cases and models occurrence rates based only on the interval over which cases occurred. The MLEs in this complete-data case are just the empirical transition frequencies.) All clusters are more likely than not to remain in a Not Contaminated state once there. A second step determines the probable number of cases conditioned on there being at least one, by sampling from the empirical probability density function (PDF) for the number of cases per quarter in each cluster. Table 3 summarizes these distributions and Table 4 shows the implied long-run expected number of cases per quarter, obtained by solving for the steady-state probability vectors of each Markov process.

Table 3: Contamination Level Probabilities

	PDF for Number of Cases per Quarter if there is at least 1			
Cluster	1 case	2 cases	3 cases	4 cases
Cluster 1	0.71	0	0.29	0
Cluster 2	0.72	0.28	0	0
Cluster 3	0.50	0.25	0.125	0.125

The distribution of the number of cases in each quarter in Australia can now be estimated via Monte Carlo simulation of the Markov processes. For each quarter, and for each institution, the current state (Contaminated or Not Contaminated) determines the next-state probabilities based on the cluster-specific state transition matrix (Table 2). If the next state is Contaminated, a second random draw determines the number of cases based on the cluster-specific probability distributions in Table 3.

Table 4: Long-Run State Probabilities

Cluster	State	
	1 Not Contaminated	2 Contaminated
Cluster 1	0.96	0.04
Cluster 2	0.85	0.15
Cluster 3	0.43	0.57

Estimating VREF Cases per Quarter in the United States

The rate of VRE infections in the USA in recent years can be estimated from data provided by the CDC National Nosocomial Infections Surveillance (NNIS) System. Data for antimicrobial resistance is tracked by the Intensive Care Antimicrobial Resistance Epidemiology (ICARE) project. Data for 1998-2000 were collected from 47 participating hospitals, with a total capacity of approximately 17,766 beds, chosen to approximate the geographic and size makeup of all US hospitals (Lawton, et al, 2000). Approximately 677 resistant cases occurred in 2000, out of 2575 isolates tested (NNIS, 2001) from intensive care units. In 2000, there were approximately 983,628 hospital beds in the US (AHA, 2001). A scale factor of 983,628/17,766 (≈ 55.36) multiplied by the number of cases yields approximate nationwide case-loads of 37,483 per year or 9,371 per quarter. Uncertainty about the annual value can be approximated by a binomial distribution with p = 677/2575 and n = 2575, or by a normal approximation to the binomial with mean of np = 677 and standard deviation of $\sqrt{np(1-p)}$ = 22.34. Thus, treating the national scale factor as fixed, the estimated quarterly VRE case-load is: (983,628/17,766)/4 × N(677, 22.34) = 13.8 × N(677, 22.34) with a mean of **9370.65 cases per quarter.**

2.2 Estimated Fraction of VRE Cases that are $VREF_A$

After estimating total VRE cases per unit time, the next step is to estimate the fraction of these cases that might involve harm to human health caused by QD treatment failures (or reduced QD treatment effectiveness) for infections with QD-resistant bacteria. For a case to fall in this subset, it must be an *E. faecium* case (rather than *E. faecalis*, which is not treated with QD.) It must also have high-level (specifically "vanA") resistance to vancomycin, since cases that are treated successfully with vancomycin are not candidates for treatment with QD. These conditions restrict the subset of cases of interest to so-called $VREF_A$ infections, meaning vanA vancomycin-resistant

E. faecium infections. After estimating the fraction of all VRE cases that are $VREF_A$ cases, we will consider how many of these cases might lead to harm from QD resistance that could potentially be prevented by banning VM use.

Estimating the Australian $VREF_A$ Fraction of VRE by Upper Bounding

Figure 2 shows all cases of VRE, including both vanA and vanB strains of both *E. faecium* and *E. faecalis*. However, almost all vanB VREF cases are susceptible to teicoplanin (Eliopoulos, 1998), so that these patients are usually not prescribed QD (Murray, 2000). Also, the JETACAR (1999) report states that, "Apart from a single strain in Australia (Butt *et al.* 1997), enterococci harboring the vanB gene have not been isolated from food or animals." Moreover, QD is not active against *Enterococcus faecali*s (manufacturer labelling, http://www.synercid.com). Thus, out of all VRE isolates, the ones that are of potential concern are the vanA VREF, i.e., $VREF_A$ cases. The count of vanA VREF in relation to other types of VRE was provided as 14/71 (0.20) in the JETACAR (1999) report and updated to 17/81 (0.21) in Turnidge (2001) with the addition of five more quarters of data. [Another study of isolates from seriously ill patients in acute care hospitals in Australia found that, of those testing positive for VRE, 6% had the vanA *E. faecium* strain (Padiglione, 2001). These isolates were not all from VRE infections, however, and could well have reflected transient colonization. Hence, to be conservative, i.e., resolving uncertainties in favour of over-estimating risk, we will use the higher fraction, 17/81 = 21%.]

To approximate uncertainty regarding the true proportion, we assume a Uniform[0, 1] (non-informative) distribution for the prior probability of a VRE isolate being a vanA VREF. This prior and binomial sampling imply a Beta(s + 1, n – s + 1) posterior, where n is the number of observations (n = 81) and s is the number of positive observations (s=17). The probability is therefore estimated as having a **Beta(18, 65)** distribution, with a mean of (s+1)/(n+2) = 18/83 ≈ **0.22.** Since this is smaller than the mean of the Uniform[0, 1] prior (i.e., 0.22 < 0.5), the choice of a uniform prior is conservative in that it tends to over-state the posterior mean. Thus, our rationale for choosing the uniform prior is not that it is strongly implied by prior knowledge or data, but rather that it acts as a plausible "upper bound" for realistic priors, i.e., it is more apt to over-state the risk from VM-related resistance than to understate it. (Of course, distributions can only be partially ordered, e.g., by whether one cumulative distribution function lies everywhere to the right of another. So, by referring to the uniform prior as an "upper bound", we only mean that it will tend to produce posterior means that lie above the true value of the uncertain fraction being estimated – here, the fraction of VRE cases that are $VREF_A$.) The purpose of biasing point

estimates and uncertainty estimates upward, i.e., so that they will be more likely to overestimate than to underestimate VM-related risk, is that we wish the final estimate to be a plausible upper bound on the true risk from continued VM use. This will allow it to be compared later (in Chapter 8) to a plausible lower bound estimate of the risk from discontinuing use.

Estimating the U.S. $VREF_A$ Fraction

According to Rice (2001), approximately 95% of VRE strains in the US are *E. faecium*. Clark *et al.* (1993) reported 82 *E. faecium* in 105 VRE isolates (78.1%) from 31 hospitals in 14 states. The State of New Jersey operated a surveillance of VRE blood isolates from 88 hospitals from 1992 to 1998 (SNJ, 2000). During that time, 70.1% of 2339 VRE samples were classified as *E. faecium*. The annual values ranged from 59.5% to 77.5%, without a clear trend up or down.

A study of 875 VREF samples was performed on human isolates from hospitals across the United States in 1994-1996 (Eliopoulos *et al.*, 1998). A subset of 352 of the total 875 isolates were identified as first isolates submitted by patients. 73% of these samples were vanA and 27% were vanB. *E. faecium* isolates submitted to the CDC from 1988 to 1992 were 83% vanA (Clark *et al.*, 1993). In a 1992 survey of 97 US laboratories, 79% of VREF isolates were vanA (Jones et al, 1995).

Using the high and low values reported for the proportion of VRE that are *E. faecium* (VREF), and for the proportion of VREF that are vanA gives a range of ($0.595 \times 0.73 = 0.43$) to ($0.95 \times 0.83 = 0.79$) for the proportion of VRE that are vanA VREF. Consistent with choosing estimates that are more likely to over-estimate rather than under-estimate risk, we can use 0.79 as a point estimate. Alternatively, for Monte Carlo uncertainty modeling, we could treat this proportion as a uniform random variable U(0.43, 0.79) with a mean of **0.61**. This specific distribution could be challenged – after all, it is plausible that further observations might uncover values outside this range – but sensitivity analysis confirms that the final results and uncertainty analysis are not very sensitive to the exact distribution chosen in this case, so long as it contains most of the range of observed data. (Indeed, the algebraic form of the risk model is a product of the 8 factors in Table 1, suggesting an asymptotic log-normal approximation to this product-of-factors. The log-normal distribution depends only on the geometric means and standard deviations of the factors, making their precise distributions unimportant as long as they are not too close to 0 or 1. We can exploit this to choose relatively simple models for parameter uncertainty, as shown in Table 1.) Finally, rather than making up any probability distribution, we can simply use sensitivity analysis to conclude that the empirical uncertainty range for

the $VREF_A$ fraction based on the above studies – from 0.43 to 0.79 – lies within a factor of 1.5 of 0.6 (i.e., (0.6/1.5) = 0.40 < 0.43 < 0.79 < 0.90 = (0.6 × 1.5), and so using 0.6 as a point estimate is unlikely to induce more than about a 1.5-fold error in the estimate. Even a 10-fold error will not change the main policy consequences of the analysis, as we shall see, so modeling the uncertainty in this parameter in greater detail is not warranted.

2.3 Fraction of Exogenous (Non-Nosocomial) VREF Cases

Most VRE are contracted nosocomially, i.e., through spread within hospitals, e.g., via unwashed hands or fomites on surfaces or medical instruments. Above, we estimated *total* VRE cases, independently of whether they were contracted nosocomially or from an exogenous source (e.g. chicken, pets, pork, etc.). However, nosocomial transmission may be viewed as being primarily a hospital-specific problem that could perhaps be eliminated by rigorous control measures but that is unlikely to be detectably affected by VM use on the farm. (See e.g., Chritiansen *et al.*, 2004 for successful control of a $VREF_B$ outbreak, and Winston *et al.*, 2002; but see also Yeh *et al.*, 2004 for the need to follow established guidelines to prevent nosocomial spread of VRE.) Hence, we will focus on exogenous (non-nosocomial) cases that are potentially attributable to exogenous sources such as chicken consumption. (However, if this restriction to cases not known to be of nosocomial origin is not desired, then the fraction estimated in this section can be replaced by 1.)

Bischoff *et al.* (1999) found that, of 347 VREF samples taken over a 5-year period at a single institution, only 31 (8.9%) were not likely to have been contracted within the hospital. Austin *et al.* (1999) developed a data-driven simulation model of nosocomial transmission dynamics of VRE in a large Chicago hospital and estimated that approximately 21% (1 out of 4.81) of cases were not due to transmission from other patients at the hospital, based on assumptions about the probabilities of transmission from HCW (health care worker) to patients and vice versa, staff-patient contact rates (patient contacts per unit time) and the average duration VRE remains transmissible on the hands of HCWs (typically one hour) and from patients (typically the duration of their stay in the ICU i.e., days). They conclude that, in the absence of stringent infection measures, 20 to 25% of cases are exogenous. A study by Thal *et al.* (1998) found 73 unique strains (via PFGE) among 379 isolates from 31 facilities in Michigan obtained between 1991 and 1996. In addition, the majority of isolates belonged to the same PFGE strains. They conclude that transmission within and between hospitals is responsible for the majority of cases. The results suggest that perhaps 73/379 or 19.3% of cases are exogenous. Although it would be useful to

have more data specifically for vanA VREF cases, we tentatively assume that vanA VREF have at least as high a proportion of nosocomial transmission as the other VRE and VREF in these studies, and then let sensitivity analysis (in which the true but unknown proportion can be as small as zero) bound the impact of this assumption on the final risk estimate.

These data suggest a plausible range of 0.089 to 0.25 for the proportion of exogenous cases. We model this via a Uniform(0.089, 0.25) distribution, with a mean of **0.170.** Again, the goal of the uniform distribution is only to approximate the range of plausible values that are consistent with the data, rather than to generate a detailed model of uncertainty. The effect of this factor is to reduce the number of cases that might be prevented by terminating VM use by a factor of about 0.17 (with an uncertainty factor of less than 2), assuming that hospitals with nosocomial transmission problems will continue to have them whether or not VM is used on the farm. Removing this assumption would therefore multiply the final risk results by about a factor of $1/0.17 \approx 6$.

2.4 Fraction of $VREF_A$ Cases Attributable to Chickens

There has been much uncertainty regarding animal sources of VREFs and more recently, QD-resistant VREFs. Willems et al (2000) used a genetic typing method [amplified-fragment length polymorphism (AFLP) analysis] to clarify potential $VREF_A$ sources by analyzing 255 VREF strains isolated from hospitalized patients, non-hospitalized persons, and various animal sources in nine different countries. Four major AFLP genogroups (groups A–D) of vanA VREFs were discriminated. Of the hospitalized patients, 4 had genogroup A strains, 10 had genogroup B strains, while the remaining 73 had genogroup C strains. *Thirty out of thirty-one chickens sampled had genogroup B strains* as did 6 of 7 turkeys. (One chicken had a genogroup C strain and 1 turkey had a genogroup A strain.) Group B strains also comprised one or more isolates from other populations, including veal calves and non-hospitalized patients. Group C strains also comprised some isolates from veal calves and all 5 isolates from cats and dogs.

While these data do not determine a precise proportion of $VREF_A$ cases attributable to chicken, they suggest that an attribution of all 10 of the genogroup B cases (out of 87 hospitalized patients) to chicken would be a generous upper bound. Again applying a Bayesian approach with a conservative non-informative (uniform) prior to quantify uncertainty regarding the true proportion gives the probability of chicken attribution as a **Beta(11, 78)** distribution, with a mean of $(s+1)/(n+2) = 11/89 \approx$ **0.12,** based on the conservative (i.e., risk-maximizing) assumption that *all* Group B

strains found in human patients are due to transfer from chickens – an admittedly extreme assumption.

If chickens seldom or never transmit resistant strains to humans, however, then the proportion of QD-resistant VREF infections that is attributable to chickens could be as low as zero. An article by Willems *et al.* (2001) supports the notion that any role of food animals in VREF infections in hospitalized patients is small. 120 epidemic and 45 non-epidemic strains of VREF isolates were obtained from hospital patients in the Netherlands, UK, US, and Australia, and 98 VREF isolates were obtained from Dutch farm animals. (Human strains were regarded as "epidemic" if they had been isolated from the same hospital, if the patients had been in contact during the outbreak period, and if the AFLP patterns showed greater than 90% similarity.) The AFLP analysis technique detected the variant esp virulence gene in 115 of 120 epidemic isolates, but in *none* of the nonepidemic isolates, and in none of the animal isolates.

In summary, genogroup data suggest that the fraction of human $VREF_A$ infections that might be due to chickens ranges from zero (or close enough to zero to be indistinguishable from it using health data) to 0.12 as a perhaps extremely conservative plausible upper bound. If it is zero, then no human health risk from vanA VREFs (QD-resistant or not) is attributable to use of VM in chickens. To obtain a plausible conservative estimate of human heath impacts, however, we will assume that *all* Group B strains found in humans are attributable to chickens (possibly after transfer to other foods or hosts through cross-contamination and secondary transmission, respectively) i.e., we will use the Beta(11, 78) distribution, with a mean of 0.12, for purposes of contingent risk modeling. All subsequent results are contingent on the validity of this assumption. In summary tables such as Table 1, the contingent nature of the analysis is indicated by explicitly listing 0 as a possible value for the chicken-attributable risk parameter.

2.5 Fraction of $VREF_A$ Cases with QD Resistance

QD Resistance Fraction Among $VREF_A$ Cases in Australia

The single best point estimate (Maximum Likelihood Estimate) of the rate of resistance to QD in VREF patients in Australia is zero, as a study of human subjects in Australia did not demonstrate any resistance to QD in human *E. faecium* isolates, despite VM being used there for many years. Specifically, a test in 1998 of 108 *E. faecium* isolates from Australian patient samples found none that was QD resistant (Turnidge and Bell, 2002). Notably, 28 of these samples were vancomycin resistant. We can incorporate this data via the conservative Bayesian approach described

above. The resistance probability has a Beta(1, 109) posterior distribution with a mean of (s+1)/(n+2) = 1/110 ≈ **0.009.** Since the value of 0 is also completely consistent with the data, it is explicitly listed in Table 1.

QD Resistance Fraction Among $VREF_A$ Cases in the United States

In the United States, $VREF_A$ incidence is relatively high and VM has also been used extensively for decades. In a study of 875 VREF samples from hospitals across the United States (Eliopoulos *et al.*, 1998), QD inhibited 98.9% of first-isolate strains at ≤ 2μg/ml, implying a 1.1% rate of intermediate or higher resistance (4 isolates). Another study utilizing 201 VREF isolates from 56 United States and Canadian medical centers in 1998 found a QD resistance rate of 1% (2 isolates resistant – none intermediate) (Jones *et al.*, 1999). Pooling the Eliopoulos and Jones data gives a total of 6 resistant isolates in 553 samples, yielding an estimated mean proportion of p **= 6/553 ≈ 0.011**, with the usual binomial uncertainty distribution.

2.6 QD Effectiveness Fraction

Any quantification of the adverse effects of QD-resistant vanA VREF must take into account that QD is not completely effective even when there is no resistance. For example, some patients cannot tolerate Synercid™, and treatment failure may occur for non-resistance reasons. In a recent study, the clinical success rate in the bacteriologically evaluable subset of patients was 70.5% with a 95% confidence interval range of 63.4% to 77.7%, corresponding to a standard deviation of 0.036 (Moellering *et al.*, 1999; Linden, 2002). Resistance to QD on therapy was observed in 6/338 (1.8%) of VREF strains. For Australia, an approximate normal probability distribution for the clinical success rate can therefore be calculated as:

Success rate = 1 – failure rate = 1 – (1 – clinical success rate – QD resistance rate))
= 1 – (1 – N(0.705, 0.036) – Beta(1,109)) = N(0.705, 0.036) + Beta(1, 109)

where N(μ, σ) denotes the normal probability distribution with mean μ and standard deviation σ. The mean value of this expression is **0.714**, close to the empirical rate of 0.705. One reason for non-success is development of QD resistance during treatment (Moellering *et al.*, 1999).

For the United States, the corresponding expression is N(0.705, 0.036) + Binomial(n = 553, p = 6/553) with a mean value of **0.716**. In summary, for both Australia and the United States, a correction factor of about 0.7 accounts for the fraction of cases for which QD treatment would be expected to succeed if QD resistance were not present. In sensitivity analyses, this

fraction can be increased up to 1 without greatly affecting the results.

2.7 Fraction of $VREF_A$ Cases Prescribed QD Over Time

$VREF_A$ cases that are not prescribed QD are not candidates for QD treatment failure due to VM-induced QD resistance. This section therefore estimates the fraction of cases that might receive QD prescriptions. This rate has been falling as new drugs (especially, linezolid) become available, and is likely to fall further in future. To illustrate how multiple time periods and forecasts of future changes can be incorporated into simple quantitative risk assessment calculations, we will use the data available in the early 2000s to forecast QD treatment rates over several quarters, and will use these time-varying estimates in subsequent risk calculations.

QD treatment is expensive, not always effective, and may have serious side effects necessitating cessation of treatment. In clinical trials, approximately 22% of subjects discontinued therapy due to adverse side effects (per product labeling). Possible alternatives for treating VREF include doxycycline and/or chloramphenicol; combinations such as teicoplanin plus gentamicin or streptomycin (Murray, 2000; Eliopoulos, 1998); linezolid (Zyvox), oritavancin, glycylcline, daptomycin, and tigicycline (Linden, 2002), while other options are in the clinical trial stage. Linezolid appears to be extremely promising and has already been approved for use in the United States, Europe, Japan, and numerous other countries.

Table 1 assumes that QD will be replaced in Australia with linezolid or other products. In time period 0 (Q4 2001), *all* VRE patients are assumed to be prescribed QD (another worst-case assumption). The decline in prescription rate is assumed to be 15% semi-annually, equivalent to 0.92 of the value in the previous quarter. This value is based on data from the United States market, where linezolid was approved in April, 2000. During Jan to June 2001, Synercid® saw a 15% decline in use while Zyvox® saw a 48% increase in use (AMR, 2001). The prescription rate can then be expressed as: 0.922^t t = 1,2,3,…. In the US, use of Synercid has declined at approximately 15% semi-annually since July 2000 (*ibid*), putting the US about six quarters ahead of Australia on the decline curve. Therefore the US prescription rate is: $0.922^{(t+6)}$ t = 1,2,3,…. when t is measured in quarters.

2.8 Time-Varying QD Resistance Fraction

The potential human health benefit (i.e., risk reduction) from banning continued use of VM in chickens arises because removing the selection pressure from continued VM use reduces the prevalence of QD resistance among $VFEF_A$ infections. But these potential benefits would not occur

immediately (and, as experience with avoparcin suggests, elimination of resistance might not occur even within decades).

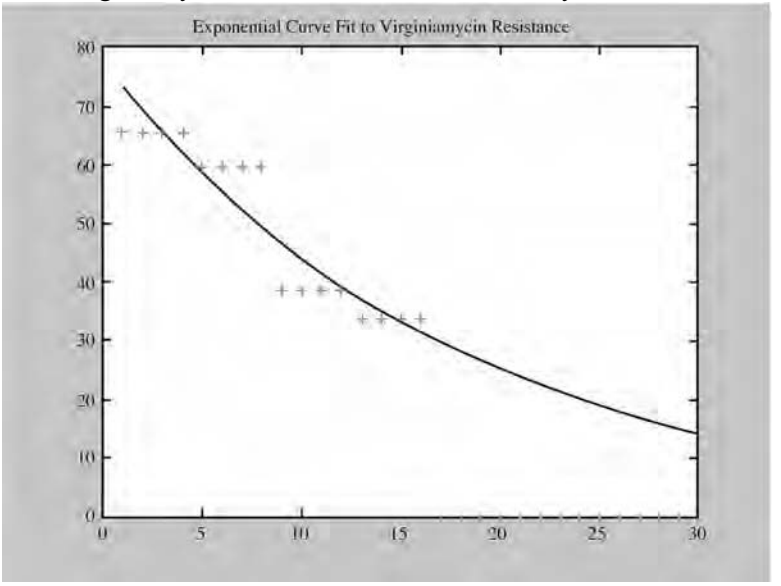
To estimate the timing of potential human health benefits of a ban on VM, it may be assumed that reductions in the QD resistance rate in $VREF_A$ in human patients are proportional to reductions in VM resistance rate in food animals (in effect assuming that QD use in animals is entirely responsible for QD-resistant VREFs in human patients.) Table 5 shows the post-ban decline in VM resistance in poultry in Denmark following the ban of VM in 1998 (DANMAP, 1997 to 2000).

Table 5: Virginiamycin Resistance in Poultry in Denmark

Year	**Resistance**	**Fraction of 1997 Level**
1997	66%	1
1998	60%	0.91
1999	39%	0.59
2000	34%	0.515

Thus after 3 years, the resistance rate is cut roughly in half. Figure 3 fits a negative exponential curve to these data.

Figure 3: Virginiamycin Resistance in Danish Poultry – Actual and Fitted



Resistance percentage as a function of time (t in quarters) is described by: Resistance(t) = $e^{(4.3526 - .0570*t)}$, so that the proportion of original QD resistance at time t is: $\mathbf{e^{(-.0570t)}}$ t = 1, 2, ….

2.9 Health Consequences of Treatment Failures

Adverse human health consequences of QD treatment failures may include excess illness-days and mortalities. Data are available from which to estimate the quantitative mortality and morbidity impacts of QD treatment failures. For mortalities, a case control study by Linden *et al.* (1997) found that 5/20 (25%) of cases given QD had VREF-associated mortality while 17/42 (40.5%) of controls receiving alternative treatment had VREF-associated mortality. This suggests a mortality rate attributable to *not* being given QD (and hence for QD resistance) of **15.5%.** This is a conservative (upwardly biased) estimate since alternative treatment options at the time of the study did not include products available now, such as linezolid. To help obtain a plausible upper bound on risk, we assume that:

Excess mortalities = 0.155 × excess treatment failures

The corresponding expected loss of life-years can be estimated from the average age and remaining life expectancy of VREF patients. A study of 262 VREF patients in the US showed a mean age of 60 years (55% females and 45% males) (Webb et al, 2001). We assume that VREF patient demographics in Australia are similar. The current life expectancy at age 60, based on insurance actuarial tables, is 79.47 for Australian males and 83.52 for Australian females (InfoChoice, 2002). Therefore, the average life-years lost per attributable mortality is:

$0.55 \times (83.52 - 60) + 0.45 \times (79.47 - 60) =$ **21.7 years**

The remaining life expectancy at age 60 is 19.6 for US males and 23.2 for US females (NCHS, 2001). Therefore, the average life-years lost per attributable mortality may be estimated as:

$0.55 \times (23.2) + 0.45 \times (19.6) =$ **21.6 years.**

To be conservative, these calculations ignore the fact that patients who receive QD treatment (e.g., leukemia or HIV patients) may have less-than-average life expectancies.

Another possible adverse effect of QD treatment failure is prolonged days of illness. A study by Linden *et al.* (2002) showed a mean duration of treatment for VREF patients treated with QD of 20 days with a range from 4 to 40 days. In an earlier study by Moellering *et al.* (1999) the mean (± S.D.) duration of treatment with QD was 14.5 ± 10.7 days (range: 1-108, n = 396). The study by Webb *et al.* (2001) found the mean days of hospitalization for

VREF patients was 48.8 days, while the mean days of hospitalization for VSEF patients was only 34.2 days. Since no analogous study for QDREF versus QDSEF patients is available – perhaps because of the lack of significant numbers of QDREF patients – we assume that additional days of treatment attributable to QD resistance is the same as for vancomycin resistance: 14.6 days. These extra days of treatment can be converted to equivalent Quality-Adjusted Life-Years (QALYs) lost if it is assumed that each extra day of illness is as bad as losing a day of life. The average number of QALY's lost due to a QD resistance attributable treatment failure (without a mortality) is then bounded by: 14.6/365 = 0.04 QALYs, as detailed in the Appendix to this chapter.

2.10 Summary

Table 6 summarizes estimated values for all model parameters for both the US and Australia.

Table 6: Summary of Model Parameters for US and Australia

Quantity	Formula (Australia; US)	Mean for Australia	Mean for US
VRE cases/quarter	Markov Simulation Model 13.84 x N(677, 22.34)	3.98	9370.65
Fixed Fractions			
$VREF_A$ fraction	Beta(18,65); Uniform(0.43,0.79)	0.22	0.61
Exogenous fraction	Uniform(0.089, 0.25)	0.17	0.17
Chicken fraction	Beta(11, 78)	0 to 0.12	0 to 0.12
QD resistance fraction	Beta(1, 109); p = 6/553	0 to 0.009	0 to .011
QD effectiveness attribution	N(0.705, 0.0362) + Beta(1,109); N(0.705, 0.0362) + Bin(6/553)	0.714	0.716
Summary of fixed reductions	Product of above fixed fractions	**0.00003**	**0.0001**
Dynamic Fractions	(t represents quarters)		
QD prescription rate	Decrease 15% semi-annually	0.922^t;	$0.922^{(t+6)}$
VM resistance fraction in chickens after ban	Decreases to 0.32 after 5 years	$e^{(-.0570\ t)}$	$e^{(-.0570\ t)}$

The factors in Table 6 are multiplied to estimate the number of VM-attributable QD treatment failures (ATF) per quarter for two alternative decision scenarios, "Ban VM" and "No-Ban VM", in the US and Australia. (Extensions to partial bans, e.g., bans on some uses but not others, are straightforward if their impacts on the factors in Table 6 are estimated. We

focus on the case of a dichotomous ban for simplicity and to obtain an upper bound for results of partial bans.) In each country, the ATF is defined as the number of QD treatment failures per quarter attributable to use of VM in chickens. It is estimated as follows. First, baseline levels are estimated by multiplying the factors in Table 6:

- ATF_{aus} = VRE cases per quarter in Australia × Beta(18, 65) × Uniform(0.089, 0.25) × Beta(11, 78) × Beta(1, 109) × [N(0.705, 0.0362) + Beta(1, 109)]. Its initial mean value is: 0.000029 × E(VRE cases per quarter in Australia) = 0.00012 cases per quarter.
- ATF_{usa} = VRE cases per quarter in US × Uniform(0.43,.79) × Uniform(0.089, 0.25) × Beta(11, 78) × Binomial(553, 6/553) × [N(0.705, 0.036) + Binomial(553, 6/553)]. Its initial mean value is: 0.0001 x E(VRE cases per quarter in US) = 0.95 cases per quarter..

Next, the time-varying component of risk is estimated by multiplying the baseline levels by factors that decline with time, as follows:

No Ban

- $ATF(t) = ATF \times$ (Proportion of patients at time t prescribed QD)
- $ATF_{aus}(t) = ATF_{aus} \times 0.922^{t}$
- $ATF_{usa}(t) = ATF_{usa} \times 0.922^{(t+6)}$.

Ban

- $ATF(t) = ATF \times$ (Proportion of patients at time t prescribed QD) × (Proportional reduction at time t in QD resistance due to VM ban)
- $ATF_{aus}(t) = ATF_{aus} \times 0.922^{t} \times e^{(-.057\,(t - T + 1))}$ $t = 1,2,\ldots 12$
- $ATF_{usa}(t) = ATF_{usa} \times 0.922^{(t+6)} \times e^{(-.057(t - T + 1)}$

where T is the quarter in which the ban takes place.

3. RESULTS

3.1 Results for Australia

Figure 4 plots results for Q1 2002 to Q4 2006 (20 quarters) from 1,000 iterations of the model for Scenario 1 (No Ban), initialized with data available as of the end of 2001. (As discussed below, these impacts would be reduced if the model were re-initialized with updated data starting in 2005, for example, since Synercid™ use has declined since 2001. But a goal of this case study is to illustrate how to use available, published data to predict future health impacts; thus, the slightly dated numbers in Tables 1 and 6 will be used for purposes of illustration.) The projected 5-year

cumulative treatment failures were 0.0011 cases, with an attributable mortality of 1.74×10^{-4} cases, representing a total of 0.004 life-years lost.

Figure 4: Cumulative Predicted Attributable Treatment Failures (Australia)

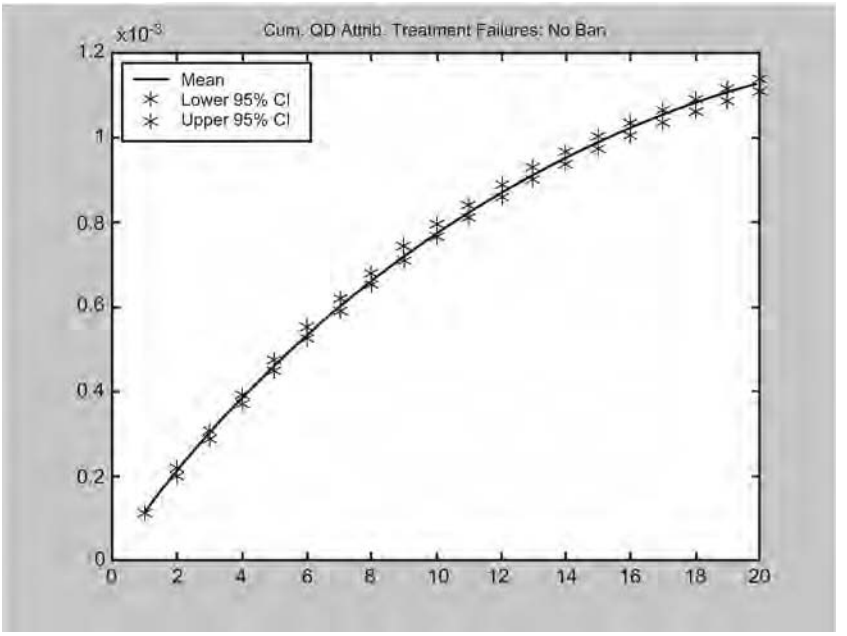


Table 7 shows results for three Banning scenarios: Ban in Q1 2002, Q1 2003, or Q1 2004; (i.e., at T = 1, 5, and 9). The differences between them stem from the fact that the later the ban, the more the switch away from prescribing QD has already occurred. The results in Table 7 imply that a ban of VM in Q1 2002 would have been likely to have reduced attributable treatment failures in Australia by at most 0.35×10^{-3} cases, mortality by 0.058×10^{-3} cases, and life-years lost by 1.3×10^{-3} over a 5-year period.

Sensitivity analyses reveal how these numbers change for variations in attribution of QD-resistant $VREF_A$ to chicken versus other sources and in the average QD prescription rate over the five year time horizon. Figure 5 shows the results of varying the chicken attribution proportion between 0 and 0.124 (its highly conservative baseline value) and varying average QD prescription rate between 0.1 and 1.0. These sensitivity analyses show a maximum of 0.001 treatment failures averted by a ban on VM, implying that at most 0.000155 mortalities in Australia would be averted in the worst case where *all* QD resistances in VREF is attributed to VM use on chickens, and all QD resistant cases are treated only with QD.

Figure 5: Range of Predicted QD Treatment Failures Averted in Five Years *vs.* Chicken Attribution Fraction and QD Prescription Rate, for Australia

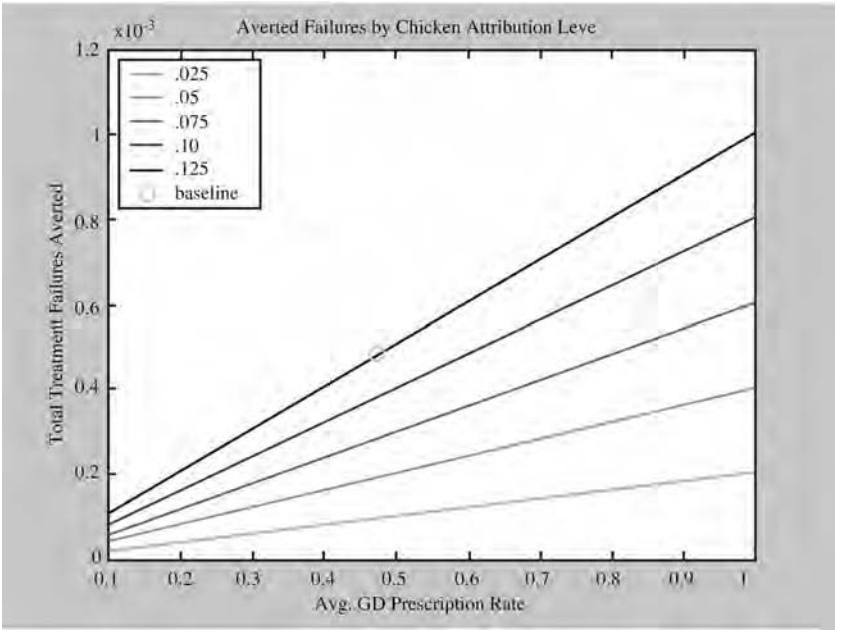


Table 7: Summary of Model-Predicted Results for Australia

	5 Year Cumulative Values x 10^{-3}		
Scenario	**Treatment Failures, 95% CI [ ]**	**Mortalities**	**QALYs**
1. No Ban	1.1, [1.1, 1.1]	0.174	3.8
2. Ban			
2a. Ban in Q1 2002 (T = 1)	0.749, [0.737, 0.761]	0.116	2.5
2b. Ban in Q1 2003 (T = 5)	0.90 [0.89, 0.92]	0.140	3.1
2c. Ban in Q1 2004 (T = 9)	1.0, [1.0, 1.0]	0.156	3.4

3.2 Results for USA

Figure 6 plots results for Q1 2002 to Q4 2006 (20 quarters) from 1000 runs of the no-ban scenario for the United States. The predicted five-year cumulative treatment failure load is 5.54 cases, with an attributable mortality of 0.86 cases, representing 18.8 life-years lost.

Figure 6: Cumulative Predicted Attributable Treatment Failures (USA)

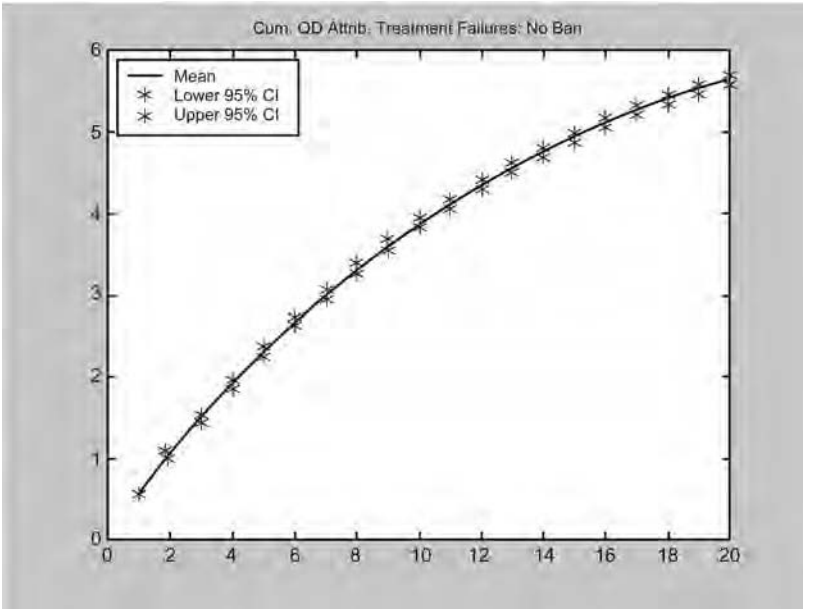


Table 8: Summary of Model-Predicted Results for United States

	5 Year Cumulative Values		
Scenario	**Treatment Failures, 95% CI**	**Mortality**	**Life-Years**
1. No Ban	5.54, [5.50, 5.58]	0.86	18.85
2. Ban			
2a. Ban in Q1 2002 (T = 1)	3.69, [3.66, 3.72]	0.57	12.56
2b. Ban in Q1 2003 (T = 5)	4.45, [4.41, 4.48]	0.69	15.14
2c. Ban in Q1 2004 (T = 9)	4.96, [4.92, 5.00]	0.77	16.88

Table 8 presents results for different ban scenarios, estimating what the effects would have been of starting VM bans in Q1 2002, Q1 2003, or Q1 2004. A ban of VM in Q1 2002 would have been predicted to have reduced attributable treatment failures by at most 1.85 cases, mortality by at most 0.29 cases (i.e., from 0.86 to 0.57), and life-years lost by at most 6.3 over a five-year period for the entire US population. Varying the chicken attribution proportion between 0 and 0.124 and varying average QD prescription rate between 0.1 and 1.0, as was done in Figure 5 for Australia, indicates a maximum of 8 treatment failures (or at most 1.24 mortalities) in the USA would be averted in the worst case where all QD resistances in VREF is attributed to VM use on chickens, and all QD resistant cases are treated only with QD.

4. DISCUSSION OF RESULTS AND KEY DRIVERS

The case study in this chapter has shown that immediately banning VM use in chickens would save less than one predicted statistical life over the next 5 years in the United States, even if it is assumed that transmission of QD-resistant *E. faecium* (or resistance determinants) from chicken to humans occurs. The effects in Australia would be far smaller, by a factor of more than 1,000 (based largely on the differences in VRE case rates per quarter and $VREF_A$ fractions in Table 6.) If no such transfer occurs, then the associated human health risk from continued use (or health benefit from a ban) is zero. In the results calculated in Section 3, a very conservative estimate of 12.4% of cases is assumed to be attributable to QD-resistance acquired from chickens, but this is probably unrealistically high. Thus, despite the scientific uncertainty about frequency of chicken-to-human transfer of resistance, it is possible to develop useful bounds on the approximate magnitude of the human health risk that might be posed by continued use of VM in chickens.

The risk estimates are driven largely by genetic typing findings showing that VREF-infected patients usually have VREF strains not found in chickens (Willems *et al.*, 2000 and 2001). Such data tightly constrain the plausible upper bound on human QD-resistance risks attributable to VM use in chickens. They make farm-to-fork modeling of microbial loads, which could be difficult or impossible with current data, unnecessary. Indeed, the key drivers of the conclusions are sufficiently strong and simple so that the Monte Carlo uncertainty analysis and sensitivity analyses presented above are unnecessary for deriving and understanding the main results. In this case study, as indicated by the tight Monte Carlo uncertainty bounds around the point estimates in Figures 4 and 6, the point estimates drive the main results.

The main drivers of the final risk estimates can be understood by considering the product of the following point estimates for key factors in the United States:

(9370.65 VRE cases/quarter) × (0.61 $VREF_A$ fraction) × (0.17 exogenous fraction) × (0.12 chicken fraction) × (0.011 QD resistance fraction) × (0.074 QD prescription rate = linezolid resistance rate).

(For simplicity, rather than using a time-varying QD prescription rate here, we simply assume that, in future, linezolid will be the first choice for treating $VREF_A$ cases and that *all* cases of linezolid treatment failure will be treated with QD. Linden *et al.*, 2002 give the linezolid failure fraction as 0.074, and we use this as the eventual future QD prescription rate. For this calculation, we ignore possible future use of daptomycin or other alternatives to QD

when linezolid treatment fails.) The product of these factors is $9370.65 \times 0.61 \times 0.17 \times 0.12 \times 0.011 \times 0.07 = 0.09$ cases per quarter, or 0.36 cases per year of QD-resistant $VREF_A$ infections treated with QD. Each such case leads to an expected fraction of a QALY lost. Even before quantifying the clinical consequences of resistance, however, it is clear that the estimated risk is limited by the relatively small fractions in this product. Including additional factors from Table 6, e.g., to reflect a gradual reduction in VM resistance rather than instantaneous elimination, reduces this product a bit further, but the main conclusion that the preventable human health harm is relatively small (probably less than 1 case of VM-induced QD treatment failure per year) can be understood on the basis of the preceding factors.

5. SUMMARY AND CONCLUSIONS

The attribution-based, product-of-factors approach to quantitative risk assessment illustrated in this chapter and summarized in Table 6 offers one possible practical template for rapidly assessing a quantitative estimate of the probable human health risk from continued use of animal antibiotics such as virginiamycin (VM). The key idea is to start with an estimate of the total rate of cases in the exposed population (i.e., the number of *all* cases of illness per unit time) and then to multiply it by a sequence of fractions to obtain a plausible upper bound on the fraction of cases that could be prevented by ceasing the animal antibiotic use. These fractions are estimated from a combination of available data and simplifying assumptions (e.g., that *all* $VREF_A$ infections having a genotype found in chickens are in fact caused by VM use in chickens, perhaps via unknown or unspecified pathways and causal mechanisms; or that withdrawing VM use would promptly eliminate *all* such cases.) These simplifications may result in an over-estimate of risk. For example, it is consistent with present microbiological knowledge that there might be no transfer of streptogramin resistance from chickens to people, so that the true risk of VM-induced QD treatment failures in people is zero. However, for purposes of guiding decision-making in the presence of important scientific uncertainties about whether a risk exists, it is appropriate to first hypothesize that it does and then to estimate its potential size, contingent on this explicitly stated hypothesis.

An advantage of arranging the calculations as a product of factors is *modularity:* each key factor can be identified, assessed, and debated separately and the effect of removing it (which is the same as the effect of dividing the entire product by that factor) can be determined immediately. For example, in Table 6, anyone who disagrees with the inclusion of an "exogenous fraction" factor of 0.17 for exogenous cases (i.e., those not

known to be of nosocomial origin) can replace it with an alternative estimate, or divide by 0.17 (i.e., replace 0.17 with 1) to remove its effects from the calculation altogether. This modularity allows attention and expertise to be focused on each factor separately. It immediately shows which factors – namely, the relatively small ones – most affect the final answer. For factors of 0.5 or more, it is unlikely that further debate and additional investigation of the exact values (or frequency distributions of values) will greatly increase the final risk estimate (since the under-estimate is at most a factor of 2 and what matters for risk assessment is usually only the order of magnitude of the risk.) Thus, the product framework also may help to focus risk assessment deliberation and debate on those factors where additional science and data can most affect the final risk estimate.

Chapter 8 will return to and expand upon the theme of modular, template-driven health risk assessment based on products of factors, but will add a crucial new component: bounding the *net* human health impacts (adverse health effects prevented minus adverse health effects caused) of alternative risk management interventions. In addition, we will consider how to extend the template-based approach to quickly account for effects of an intervention (such as ban of an existing animal antibiotic use or introduction of a new one) on multiple pathogens in multiple food animals. However, before further developing this product-of-factors approach, in which each factor is estimated from available data and simplifying assumptions, we will first examine in depth the extent to which current and historical data can be relied on to estimate likely future health consequences of interventions. This requires uncertainty analysis of a predictive systems dynamics model that considers the dynamics of resistance emergence under selection pressures from both animal and human antibiotic use.

Appendix: Estimating Human Health Consequences of QD Resistance

VM use in animals can adversely affect human health only if there is a difference in the human health consequences of QD-resistant vs. QD-susceptible strains of VanA VREF, perhaps due to differences in the medical treatment received. This appendix estimates this difference.

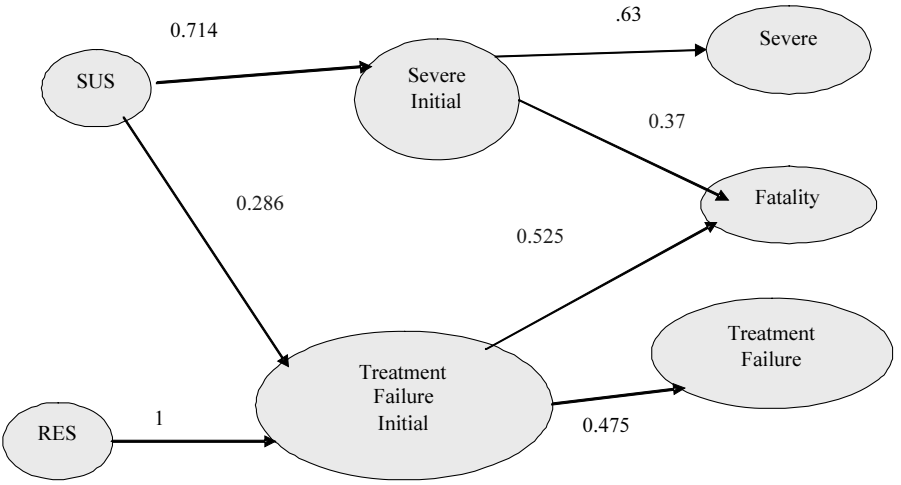
Figure A1 illustrates possible health outcomes for QD-susceptible (SUS) and QD-resistant (RES) vanA VREF cases, with their estimated probabilities. Each VanA VREF patient has an initial health state classified as either "severe illness" (but susceptible to QD treatment) or as "QD treatment failure" (= not susceptible to effective QD treatment). Each has a final health outcome of either "Severe" illness (but responded to QD treatment), QD "Treatment failure" (but no fatality), or "Fatality". An initially severely ill vanA VREF-infected patient who does not respond to QD enters the "Treatment Failure" state, from which progression to the "Fatality" state may then occur. Transition from Severe illness to the Treatment Failure state can occur not only if the vanA VREF infection is QD-resistant, but also for reasons such as a patient's intolerance of QD. (These latter could be classified as "unable to treat with QD" rather than as "treatment failures", but the distinction is not needed for quantifying the increase in treatment failures due to resistance.) Transitions to the "Fatality" category can occur from both the severe illness and treatment failure initial categories, but the probability of fatality is higher for treatment failures (0.525 vs. 0.37).

The transition probabilities from Severe Illness to Treatment Failure and from both Severe Illness and Treatment Failure to Fatality are estimated as follows. Lautenbach et. al. (1999) reported a 37% VRE-attributable mortality rate among patients with enterococcal bacteremia, 28% of whom were resistant to vancomycin. A case-control study by Edmond, *et al.* (1996) found a mortality rate of 37% (95% CI = [0.1, 0.64]) for vancomycin-resistance among cases of enterococcal bacteremia, gives an estimated value of:

$$\Pr(\text{fatality} \mid \text{severe illness and no QD resistance}) = \mathbf{0.37}$$

Linden *et al.* (1997) found that 5/20 (25%) of cases given QD had VREF-associated mortality while 17/42 (40.5%) controls receiving alternative treatment had VREF-associated mortality. These proportions are borderline statistically significantly different ($p = 0.067$ in a two-sided test for difference of proportions, $p = 0.034$ for a one-sided test). Their difference of $40.5 - 25 =$ **15.5%** can be used as a plausible conservative estimate of the increase in mortality probability attributable to not being treated successfully with QD. (If the true difference is zero, the total risk associated with QD resistance would also be zero.) The estimated fatality probability among Treatment Failure patients is thus:

Figure A1: Outcome Probabilities for QD-Resistant (RES) and QD-Susceptible (SUS) VREF Cases



Pr(fatality | QD treatment failure) = 0.37 + 0.155 = **0.525**

The probability of treatment failure is defined as 1 among QD-resistant cases:

Pr(QD treatment failure | QD resistance) = 1

However, QD is often not completely effective, even in the absence of resistance. In a recent study, the clinical success rate in the bacteriologically evaluable subset of patients was 70.5% with a 95% confidence interval range of 63.4% to 77.7% (Moellering *et al.*, 1999; Linden, 2002). A fraction of the treatment failures (about 0.009, Cox and Popken, 2004) were due to resistance and must be subtracted out to get the fraction of failures that are not due to QD- resistance:

Pr(QD treatment failure | QD-susceptible) = 1 - 0**.**705 - 0.009 = 0.2860

Multiplying (1 – 0.2860) = 0.714 by the 15.5% estimated increase in mortality rate among cases not treated successfully with QD (for any reason) gives **0.11** as an estimate of the increase in mortality rate due specifically to QD resistance-related treatment failures (rather than other causes of treatment failure.)

The final outcome probabilities for QD-resistant and susceptible infections are computed as follows:

- Pr(severe illness is the final outcome | QD-resistant) = **0** (see Figure A1).

- Pr(fatality | QD-resistant) = Pr(fatality | QD treatment failure) = **0.525**
- Pr(treatment failure but no fatality | QD-resistant) = 1 – 0.525 = **0.475**
- Pr(severe illness is the final outcome | QD-susceptible) = Pr(severe illness & successful treatment & no fatality | susceptible) = 1*(1 – 0.286)*(1 – 0.37) = **0.45**
- Pr(QD treatment failure but no fatality is final outcome | QD-susceptible) = (0.286)*(1 – 0.525) = 0.286*0.475 = **0.136**
- Pr(fatality is final outcome | QD-susceptible) = Pr(fatality | severe initial outcome) * Pr(severe initial outcome | QD-susceptible) + Pr(fatality | treatment failure) * Pr(treatment failure | QD-susceptible) = 0.37*(1 - 0.286) + 0.286* 0.525 **= 0.4143**
  - Pr(fatality | severe initial outcome & QD-susceptible) = 0.37
  - Pr(severe initial outcome | QD-susceptible) = 1 – 0.286 = 0.714
  - Pr(QD treatment failure | QD-susceptible) = 0.286
  - Pr(fatality | QD treatment failure) = 0.525 (see Figure A1).

Severity of Consequences: Estimating QALY Losses

VREF patients in a recent study (Webb *et al.* 2001) incurred an average of 14.6 additional days of hospitalization compared to VSEF patients (34.2 days *vs.* 48.8 days). No analogous study for QDREF versus QDSEF patients is available, perhaps because of the small numbers of QDREF patients. We will assume that the additional number of days of treatment attributable to QD resistance or treatment failure is the same as the additional days attributable to vancomycin resistance, i.e., **14.6 days.**

Extra days of treatment and illness can be converted to lost quality-adjusted life-years (QALYs). The HUI3 multiattribute utility scale (Furlong et al, 2001) provides values from –1.371 to 1.00, with negative scores representing states considered to be worse than death. Other scoring systems provide values from 0.0 to 1.0. HUI3 requires rating patients in 8 health attributes with scores ranging from 1 to 6 and converting the results to a single value via Multi-Attribute Utility Theory (Cox, 2001). For simplicity, we will use a conservative overall rating of 0.0 for VREF patients with QD resistance during their treatment. The average numbers of QALYs lost per case are then as follows:

- Severe (susceptible): 48.8/365 = 0.1337
- Treatment Failure (susceptible or resistant): 63.4/365 = 0.1737

Thus, each treatment failure (e.g., due to QD-resistance) is expected to generate an additional **0.04** QALYs lost per non-fatal treatment failure. This is a conservative estimate, as it *assumes that no other effective therapies are applied* following QD failure.

The QALY's lost per fatality are estimated by first comparing the average age of a VREF patient to average life expectancy. A study of 262 VREF patients in the

US determined a mean age of 60 years, composed of 55% females and 45% males (Webb *et al.* 2001). The current life expectancy at age 60, based on insurance actuarial tables, is 79.47 for males and 83.52 for females (InfoChoice, 2002. We do not calculate separate values for the US and Australia for purposes of this rough estimate.) Therefore, the estimated average QALYs lost per attributable fatality is: 0.55*(83.52-60) + 0.45*(79.47-60) = **21.70 years.** This number is conservative (risk-maximizing), in that it assumes that a seriously ill VREF patient would have the same life expectancy and QALYs as a member of the general population if QD therapy were effective.

Chapter 7

Dynamic Modeling and Uncertainty Analysis

1. INTRODUCTION

Recent qualitative analyses have warned of potential future human health risks from emerging resistance to antibiotics in food-borne pathogens due to use of similar antibiotics in both food animals and human medicine. While historical data suggest that human health risks from some animal antimicrobials, such as virginiamycin (VM), have remained low (McDonald *et al.*, 2001), there is widespread concern that "resistance epidemics" or endemics could arise in future. Systems dynamics models of exposure and risk, introduced in Chapters 4 and 5, can not only help to quantify the likely transient responses of illness and resistance rates among human and animal patients in the years immediately following a change in antibiotic use, but can also inform decision-makers about the risks of longer-run emergence of resistance and the effects of different antibiotic use patterns on hastening or slowing it.

This chapter examines a recent deterministic systems dynamics model of the emergence of resistance to antimicrobial drugs in human patients. As a case study, we continue the example of emergence of streptogramin resistance in vancomycin-resistant *E. faecium* (VREF) in response to continued use of VM in chickens. The original version of the model, taken from the recent literature (Smith *et al.*, 2002, 2003), provides deterministic estimates of future human health risks from continued VM use. This chapter generalizes the approach by allowing for uncertainty in the model parameters and applying a Bayesian framework to construct refined parameter estimates that are consistent with historical data. We also extend

the deterministic mathematical model (Smith *et al.*, 2002) to more accurately represent the stochastic dynamics of resistance spread when resistance prevalence is rare. The resulting stochastic simulation model is applied to data for the human streptogramin combination antibiotic quinupristin-dalfopristin (QD) to obtain quantitative bounds on the future human health risks to QD patients from continued use of VM in food animals.

Although this chapter is more specialized and technical than the rest, the Bayesian approach to uncertainty analysis of dynamic models that it illustrates may be useful in future as systems dynamics models become increasingly used in microbial and antimicrobial risk assessments.

1.1 Motivation: How Does History Constrain Future Risks?

A frequent conundrum in applied risk analysis is how to interpret negative data. How reassuring about the future is the non-occurrence of a feared event to date, despite prolonged exposure to a hypothesized hazard? For example, if a hazardous facility has not produced a catastrophic accident after many years of operation, how strong is the evidence that the true risk per facility-year is not very large? If the last N items sampled from a production process have zero defects, how much confidence should this give that the true defect rate is not greater than some specified acceptable level? Or, turning to the main topic of this chapter, if decades of use of an animal antibiotic such as macrolides or VM have not yet led to significant human resistance to similar human antibiotics, then how strongly (if at all) does this justify an inference that large-scale human antibiotic resistance is unlikely to emerge in future if current use continues?

Simple methods suffice to interpret negative data in simple situations. For example, the "rule of three" specifies $3/N$ as an approximate non-parametric upper 95% confidence limit for the true proportion of defects if 0 defects have been found among N randomly sampled items. A tighter bound sometimes results by applying Bayes' Rule. If s failures are observed out of N cases examined, each with the same true but unknown probability of failure, then the expected value of the true failure probability starting from a flat (uniform) prior after conditioning on these data is the mean of a Beta posterior distribution, given by $(s + 1)/(N + 2)$, or by $1/(N + 2)$ if $s = 0$. In these cases, where the true risk remains constant over time, the estimated risk after N observations of no defect is of order $1/N$. More generally, for special families of systems, such as decreasing failure rate (DFR) reliability systems or probabilistic mixtures of such systems, observed histories provide useful upper bounds on potential future risks (Ross, 1996).

But no guarantees about future hazard rates can hold in the most general case, when the underlying probability of failure may increase

randomly over time. As a counterexample, suppose that a system deteriorates by making unobserved transitions through a sequence of states, $1 \rightarrow 2 \rightarrow 3 \rightarrow \ldots$ with the hazard rate for system failure from each state being uncertain and greater than in previous states. Then a long history of no failures may simply indicate that the system is (probably) still in its early, low-risk, states, with higher failure rates still likely to emerge at some random time in the future when the system makes a transition to a more hazardous state. In this case, Bayesian inference may be used to update prior probabilities for the current state of the system and hence for current and future risks, to reflect the passage of time. But bounds on future risks cannot be derived from past data alone, since they depend on the hazard rates of as-yet unobserved future states of the system. Predicted risks are then sensitive to modeling input assumptions about the values of future state-specific hazard rates.

In many practical settings, however, the situation is less extreme. The future state of a dynamically evolving system may be uncertain because of stochastic evolution and because of uncertainty about the true values of transition parameters, but the possible states are known and at least some transitions have been observed from which plausible bounds on the transition rates can be estimated. This chapter deals with such a situation of current practical interest in which model parameters (e.g., the transition rates among states) remain constant over time, but population risks may evolve and randomly increase as members of the exposed human population make stochastic transitions among different states, such as unexposed, exposed-and-infected, infected-and-contagious, infected-and-hospitalized, and recovered. Historical data from which to estimate state-specific transition rates are available, but the question of what they imply for potential future illness rates in the population remains to be answered. This question arises in the context of an important public health risk management decision: whether banning animal antimicrobial use (AAU) among food animals on farms will significantly protect human health by decreasing the rate at which resistance to analogous human antimicrobial drugs emerges in patient populations. We address the question of how to use past data to put bounds on plausible future risks to human health from AAU in this context.

1.2 A Growing Concern: Managing Potential Future Risks

Emergence of resistance to human antibiotics among food-borne bacteria from food animals exposed to similar or identical drugs poses human health risks that are currently highly uncertain. Yet, it is intuitively plausible that resistance may increase stochastically over time, driven in part by selection pressures from antimicrobial use on farms. If food-borne bacteria that contribute to human illnesses increase treatment failures when

human patients are treated with antimicrobials similar to those used in animals, or if animal antimicrobial use (AAU) selects for bacterial strains that exhibit cross-resistance or co-resistance between animal and human drugs, then AAU-induced selection pressures may result in increased average illness-days per capita-year in the human population or in certain subpopulations such as infants, the elderly, or patients with compromised immune systems. Quantifying the potential for increase in human health risks caused by AAU requires modeling the stochastic dynamic relation between AAU, selection pressure, and emergence of resistance in human populations and among patients with food-borne bacterial illnesses.

Smith *et al.* (2002) propose one such model. Their simple theoretical mathematical analysis, reviewed below, implies that selection pressures from combined AAU and medical antibiotic use (MAU) in people can speed development of antibiotic resistance (AR) in human commensal bacteria. Their modeling leads to the dramatic conclusion that:

> "Small increases in prevalence when AR is rare have dramatic effects, like sparks that start forest fires. … Our analysis suggests that AAU hastens the appearance of AR in bacteria in humans. Our model indicates that the greatest impact occurs very early in the emergence of resistance, when AR bacteria are rare, possibly below the detection limits of current surveillance methods. … Regulating early AAU would likely extend the period that a drug can be used effectively in humans" (Smith *et al.*, 2002).

This model provides both an explicit framework for understanding and simulating potential human health risks from AAU and also a plausible explanation for why data supporting the hypothesis that such risks might exist might not be readily available: because the increase in risk occurs while AR bacteria are still too rare to be easily detected.

A subsequent paper (Smith *et al.*, 2003) focused specifically on how use of virginiamycin (VM) in animals and of the nearly identical streptogramin combination quinupristin-dalfopristin (QD) in people might increase human health risks. QD was approved in 1999 to treat vancomycin-resistant *Enterococcus faecium* (VREF) bacterial infections in human patients. As discussed in Chapter 6, such patients typically have severely compromised immune systems that cannot defend successfully against infection by the *E. faecium* commensal bacteria normally found in human intestines. Smith *et al.* concluded that:

> "Virginiamycin, another streptogramin, threatens the efficacy of QD in medicine because streptogramin resistance in enterococci associated with food animals may be transferred to *E. faecium* in hospitalized patients…. We

> conclude that the emergence of SREF [i.e., streptogramin-resistant VREF] is likely to be the result of an interaction between QD use in medicine and the long-term use of virginiamycin for animal growth promotion. Virginiamycin use has created a credible threat to the efficacy of QD by increasing the mobility and frequency of high-level resistance genes."

These model-based conclusions were promptly seized on by anti-animal antibiotic activist groups to urge that "FDA has an obligation to regulate virginiamycin because there's a reasonable expectation that its continued use in animals will accelerate the evolution of Synercid-resistant bacteria" (Keep Antibiotics Working, 2002). They resonated with fears and speculations about human health risks from AAU stoked in popular books and some journal articles (APUA, 2002) through rhetoric such as: "But given how easily enterococci appeared to pass resistance genes to one another as a general matter and how easily VRE circulated in hospitals, how long would it be before Synercid-resistant VRE was ubiquitous? Before the new miracle drug was dead?" (Shnayerson and Plotkin, 2002, p. 119).

Smith *et al.*'s clear caveats that their conclusions follow from unvalidated, possibly incorrect, modeling assumptions, were ignored in headlines that reported only that "The use of antibiotics to promote growth in farm animals hastens the end of their medical effectiveness" (http://www.sciencenews.org/20020427/fob1ref.asp).

1.3 From Qualitative Hazard Identification to Quantitative Risk Assessment

To inform rational risk management decision-making, it is useful to extend the mathematical modeling work to quantify the magnitude of the human health risks from AAU, using virginiamycin and Synercid™ (QD) as a case study. *How large* is the threat to QD efficacy for human patients and *by how much* does use of VM in animals hasten the appearance of QD-resistance in human *E. faecium?* What is the numerical answer to the above rhetorical question "How long would it be before Synercid-resistant VRE was ubiquitous?" Answers to such quantitative questions are crucial for risk managers. They are addressed in the following sections by extending the Smith *et al.* mathematical modeling framework to support discrete-event simulation (DES) (rather than the continuous ordinary differential equation (ODE) approximation of Smith *et al.*), to capture more accurately the stochastic dynamics of AR spread when AR prevalence is rare. We apply the resulting stochastic simulation model to real data for VM and QD to obtain quantitative bounds on the future human health risks to QD patients from continued VM use in animals.

2. DETERMINISTIC MODELING AND RESULTS

This section summarizes key aspects of the mathematical model of Smith *et al.* (2002) and discusses its re-implementation as a discrete-event stochastic simulation model. The purpose is to understand and predict effects of AAU use and other parameters on the timing and magnitude of adverse human health effects, specifically, cases of Synercid™ treatment failure or resistance.

2.1 VRE Colonization Dynamics Model

Figure 1 diagrams the major flows in the model, as individuals make transitions among the four health states of "unexposed", "exposed", "colonized" (with resistant bacteria), and "amplified". The "amplified" state refers to a condition of being colonized and highly contagious. The fractions of the population in these four states are denoted by the following four variables:

- W(t) = fraction of population unexposed at time t
- X(t) = fraction exposed at t
- Y(t) = fraction colonized at t
- Z(t) = fraction amplified at t

Prevalence at time t is defined, following Smith *et al.*, as X(t) + Y(t) + Z(t). A critical *coefficient of secondary transmission,* R_0, gives the average number of humans exposed by each exposed human when prevalence is approximately zero. $R_0 = 1$ represents an epidemic threshold above which a new resistant strain spreads and persists.

Transitions of individuals among the four compartments, W, X, Y, and Z occur at average rates (per person-day) determined by the transition rate parameters in Table 1. The right-most column of Table 1 lists the values of these parameters for VRE dynamics used by Smith *et al.* for purposes of illustration**.**

Figure 1: Diagram of discrete-event simulation (DES) model

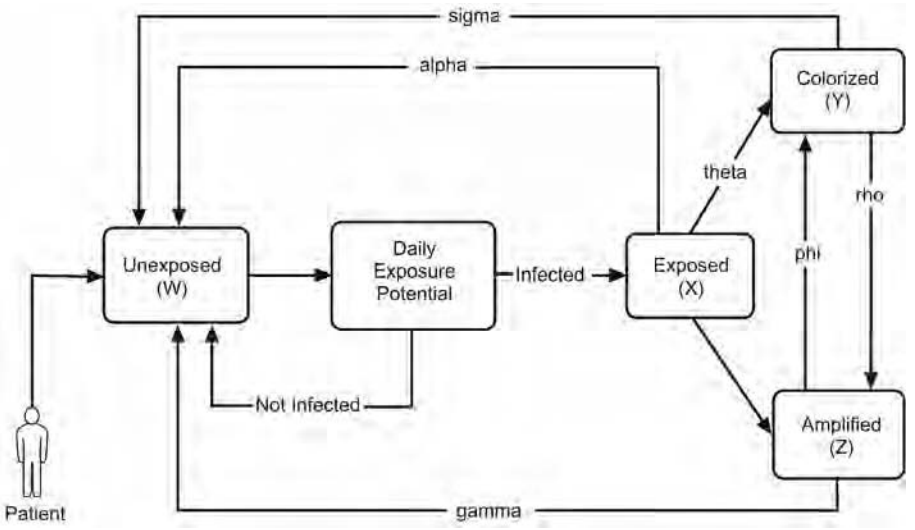


Table 1: Input parameters for VRE dynamics model

Symbol	Interpretation	VRE value
λ	rate of exposure per capita due to animal antibiotic use (AAU rate)	0.001
μ	background rate of exposure per capita due to sources other than AAU	10^{-6}
α	transient loss rate/day = transition rate from exposed to unexposed, per exposed person per day	0.1
θ	colonization rate/day = transition rate from exposed to colonized, per exposed person per day	0.001
σ	colonized loss rate/day = transition rate from colonized to unexposed, per colonized person per day	0.003
γ	amplified loss rate/day = transition rate from amplified to unexposed, per amplified person per day	0.007
φ	recolonization rate/day = transition rate from amplified to colonized (per amplified person per day)	0.003
ρ	prescription rate/day (MAU rate)	0.003
η	contact rate (colonized) /day	10^{-5}
β	contact rate (amplified) / day	0.5

Source: Smith *et al.,* 2002

Smith *et al.* approximate the stochastic transition system just described via the following system of nonlinear deterministic ODEs and algebraic equations for the mean values of its variables:

Summary of Deterministic Model Equations

$dX(t)/dt = [\mu + \lambda + \eta\, Y(t) + \beta\, Z(t)]W(t) - (\theta + \alpha + \rho)X(t)$ = rate of change in exposed population fraction a time t

$dY(t)/dt = \theta X(t) + \varphi\, Z(t) - (\rho + \sigma)Y(t)$ = rate of change in colonized population fraction

$dZ(t)/dt = \rho[X(t) + Y(t)] - (\gamma + \varphi)Z(t)$ = rate of change in amplified population fraction

$W(t) = 1 - [X(t) + Y(t) + Z(t)]$ = rate of change in unexposed population fraction at time t

$EY = (\sigma + \rho)(\theta\gamma + \theta\varphi + \rho\varphi) / (\rho + \alpha + \theta)(\gamma\sigma + \gamma\rho + \varphi\sigma)$

$EZ = \rho(\gamma + \varphi)(\sigma + \rho + \theta) / (\rho + \alpha + \theta)(\gamma\sigma + \gamma\rho + \varphi\sigma))$

$R_0 = EY\eta/(\rho + \sigma) + EZ\beta/(\gamma + \varphi)$

where EY and EZ are the expected number of times that a single exposed individual is colonized or amplified, respectively.

The value for R_0 determined by these equations and by the input parameters in Table 1 is 1.98, implying that resistance will increase with certainty. (Recall that $R_0 = 1$ is the epidemic threshold for resistant strains in this model.) Even starting from a prevalence of 0, the input values in Table 1 result in an equilibrium population prevalence of approximately 50% after about 5 years – a level about 50 times higher than the current resistance rate (McDonald *et al.*, 2001). Thus, for the assumed input values in the right-most column of Table 1, a high persistent (endemic) level of resistance is predicted to emerge within just a few years, apparently justifying the worst fears of those who urge the immediate ban of virginiamycin in order to keep Synercid™ effective.

Sensitivity Analysis Results for the Deterministic Model

The prediction of rapidly emerging human Synercid™-resistant bacteria (SREFs) is very sensitive to the assumed input values in Table 1. Barber *et al.* (2003) promptly noted that:

> "There are generally no cited references or defenses for the values used in the above model, and sensitivity analysis of these parameters reveals some interesting findings… The model is particularly sensitive to the parameters (α, β, and γ). Examining first the level of α, assume that instead of 0.1 per day of transient loss, we use α = 0.38. This would be equivalent to a little more than one third of the exposed individuals simply passing the resistant bacteria through their system. With this level of α, equilibrium prevalence falls to about 0.01, compared to about 0.50, which is obtained using their stated assumptions [in Table 1 above]. …Similarly, the assumed degree of exposure from amplified individuals is quite high, with β = 0.5 per day. A value of β = 0.1 per day drives the equilibrium prevalence to less than 0.03. … A γ = 0.07 results in an equilibrium prevalence less than 0.02." (Barber *et al.* 2003)

These deterministic sensitivity analyses of Barber *et al.* establish that varying the input values assumed in Table 1 can significantly change the predicted level of resistance in the human population. But they do not address the questions of which parameters must be modified (if any) to be consistent with available data, nor of whether the values discussed by Barber *et al.* are more likely to be correct than the ones in Table 1. To address these questions, it is useful to examine the consistency of different parameter value combinations with available data using Bayesian uncertainty analysis. As shown below, more realistic input values based on data for streptogramin-resistant VREFs (SREFs) lead to substantial revision of some parameter values, including for λ and also for the prescription rate, ρ (not discussed by Barber *et al.*) Re-estimating these parameters based on QD-specific data leads to very different conclusions from those implied by the parameter values in Table 1, as deterministic sensitivity analysis suggests could happen.

3. BAYESIAN MONTE CARLO (BMC) UNCERTAINTY ANALYSIS

To obtain refined estimates of the parameters in Table 1, it is natural to consider applying Bayesian inference to incorporate relevant historical information on observed QD resistance rates. To this end, at least some of

the uncertainty about the parameters in Table 1 must first be described by prior probability distributions that allow for the possibility that the estimated values are not exactly correct. Next, historical data on observed levels of VM use and resistance in human populations are required. Any realistic model should be consistent with these data. Finally, Monte Carlo simulation with rejection of samples that are inconsistent with available data (i.e., samples that predict that very high resistance rates would almost certainly already have emerged, in contradiction to historical experience to date) is used to derive posterior distributions that are conditioned on (i.e., consistent with) the observed data. Bayesian Monte Carlo uncertainty analysis based on rejection of sampled parameter value combinations that are inconsistent with observed data, sometimes called Probabilistic Logic Sampling (PLS), has been well-developed in the AI-and-Uncertainty and computational statistics literatures (e.g., Henrion, 1988). While more sophisticated extensions such as adaptive importance sampling are now available to reduce CPU time, the current application is simple enough so that PLS is directly applicable.

To allow accurate quantification of the impact of continued VM use on rare-event health outcomes, such as the frequency of small outbreaks of SREFs in hospital wards, we also extend the model in Figure 1 to a stochastic discrete-event simulation version by reinterpreting the deterministic flow rate parameters as the intensities of stochastic transition rates (Poisson intensities), with units of expected transitions per unit time among compartments (disease states). The PLS Bayesian Monte Carlo uncertainty analysis applies equally well to either the deterministic or the stochastic formulations, but the stochastic formulation gives information on probability distributions and rare events as well as expected values of compartment sizes over time.

3.1 Selecting Priors for Probabilistic Sensitivity Analysis

Selection of prior distributions can be contentious in applied Bayesian risk assessment, with no truly satisfactory objective resolution. To avoid pretending greater certainty than we have about what priors should be used, we took the initial values in Table 1 as estimates of the prior means and created distributions with fairly modest uncertainties around them. Specifically, we chose log-normal distributions for the parameters λ, μ, θ, φ, ρ η, with their small values and high uncertainties, and normal distributions for the other four parameters in Table 1. The estimated values in Table 1 were taken as prior means and were assumed to have a 95% probability of lying within 1 $\log_{10}$ (for λ, μ, θ, φ, ρ η) or within 50% (for the remaining parameters) of their true values. Then, we interpreted these priors not as best

guesses at true values, as is often done, but rather as conservative (i.e., overly narrow) expressions of uncertainty about the estimated values in Table 1. We concede (and even expect) that the true values of some parameters could well be outside the plausible range of values specified by our priors, but in this case, Bayesian inference should place the posterior means for such parameters in the tails of their prior distributions, indicating that perhaps a less conservative prior (i.e., one that allowed for a larger range of plausible values) would have been justified and that, in any case, there is reason to revisit the value assumed for that parameter in Table 1. We thus treat the Bayesian inference process as a form of *probabilistic sensitivity analysis* in which available data are given the opportunity to pull posterior values far from their prior means or to leave them relatively close (e.g., within a small number of standard deviations), depending on how well they fit the data. With this interpretation, the Bayesian analysis becomes a diagnostic tool capable of showing whether available data require any large adjustments to initial estimates of the model parameters.

3.2 Data on Historical QD Resistance Rates

VM has been heavily used in poultry in the US since 1975, while QD has been in use in humans since 1999. However, high QD resistance has not yet been observed in the US population (McDonald *et al.*, 2001). For example, prior to its approval for humans, a study of VREF samples on human isolates from hospitals across the United States in 1994-1996 (Eliopoulos *et al.*, 1998) indicated a 1.1% rate of intermediate or higher resistance (4 isolates). Another study utilizing 201 VREF isolates from 56 US and Canadian medical centers in 1998 found 2 that were QD resistant (Jones, *et al.*, 1999). A study by Linden, *et al.* (1997) found 5 cases of emerging (*in vitro*) QD resistance among 396 VREF patients given QD on an emergency use basis. McDonald *et al.*, (2001), in a 1998-1999 study, found that 3/334 *E. faecium* isolates were QD resistant. From these studies, a pooled estimate of preapproval QD resistance rate of approximately 14/1283 = 1.1% (range of 0.9% to 1.26%) is plausible. (In general, pooling data from diverse studies using different microbiological methods and endpoints can raise difficulties of interpretation and requires caution, but in the present case, the different numbers are close and are used only to suggest a historical rate on the order of 0.9% to 1.26%.) Therefore, any more recent study finding QD resistant VRE at levels much higher than 1.26% would lead one to suspect that QD resistance is becoming more prevalent. For analysis purposes, a *detection threshold*, d, representing the true prevalence of QD resistance at which detection of significant increases from preapproval rates is highly likely, can be used in sensitivity analyses.

3.3 Bayesian Monte Carlo (BMC) Estimation of Posterior Distributions

To obtain posterior distributions for the parameters, PLS Bayesian Monte Carlo uncertainty analysis (Henrion, 1988) was applied to the discrete-event stochastic simulation version of the model in Figure 1, i.e., the simulation model was run many times with input values randomly sampled from their estimated prior distributions and runs that were consistent with observations (i.e., the observed low resistance rates) were kept. Each run (iteration) of the simulation begins in 1975, when VM was introduced in poultry; continues through 1999, when human use of QD began; and ends at year-end 2002 (usage dates from McDonald *et al.*, 2001). The human prescription rate for QD (Synercid™) is set to zero until 1999, the year it was approved for human use. Each iteration of the simulation draws each parameter from its respective prior distribution. The simulation is then run up to year-end 2002. If the population prevalence of QD resistance at that time is below the inconsistency-detection threshold *d* (i.e., the prevalence is consistent with observations), then the parameter values from that iteration are recorded. Iterating many times generates a joint frequency distribution of parameter value combinations that are consistent with the historical lack of high QD resistance (using detection threshold *d*). These values are used to form posterior and predictive distributions for the parameters and for the corresponding probability that the coefficient of secondary transmission exceeds the epidemic threshold value, $R_0 > 1$. [To speed calculations, posterior distributions within simulation loops were estimated via Bayesian conjugate analysis for normal distributions (Sebastiani, 2001).]

Stochastic Discrete-Event Simulation (DES) of Time-to-Outbreak

The stochastic DES version of the model in Figure 1 can simulate the probability distribution of the time until an "outbreak" (defined here as two or more simultaneous cases of patients in the "amplified" state) of resistant bacteria occurs. Because of random transients, an outbreak can occasionally occur even if $R_0 < 1$. Our DES model simulates a hospital intensive care unit (ICU) as a fixed-size (e.g., 700-bed, following Smith *et al.*, 2003) population of patients makes transitions among the health states *unexposed, exposed, colonized*, and *amplified*. The simulation entities are individual patients. Hospital/community exchange dynamics need not be explicitly modeled; instead, for simplicity, unexposed patients who leave (i.e., return to the community) are assumed to be replaced with new unexposed patients

(assuming that the ICU remains full). Amplified cases are assumed to occur in the ICU.

In accord with the theory of Markov chains, the parameter estimates in Table 1 are used to estimate the branching probabilities from each current state to each potential next state in Figure 1: they are just the ratios of the next-state-specific transition rates to the total (sum over all potential next states) transition rates from each state. The random time spent in each state with exposure is then given by an exponential distribution whose mean is the reciprocal of the total departure rate from the state. For the *unexposed* state, the departure rate changes as a function of the number of patients in the *colonized* and *amplified* states. In this case, simulated patients are subjected to an exposure potential each simulated day. The exposure potential is the sum of the exposure rates: daily exposure rate = $\lambda + \mu + \eta Y + \beta Z$

Running many iterations of the simulation model determines the probability of an outbreak occurring in any year and the probability distribution for the time until an outbreak of streptogramin resistant *E. faecium* (SREF) occurs.

4. RESULTS OF UNCERTAIN DYNAMIC MODEL

4.1 Initial Bayesian Analysis and Estimation of Parameters

To compute initial posterior estimates of the parameters in Table 1 consistent with historical data, 10,000 simulation iterations were first run using Smith's example parameter values from Table 1 as the prior means and using an assumed detection threshold of d = 1.5% to define potentially significant elevations above the historical rate of 1.1%. 920 of these simulations produced a prevalence lower than d = 0.015. The corresponding simulated parameter value combinations were used to form the posterior distributions, as described above. Table 2 shows the prior and posterior means for the model input parameters based on these initial runs.

One feature of the initial runs was that the posterior mean for the prescription rate, ρ, for QD to patients was significantly less than the prior value estimated in Table 1 (0.0008 vs. 0.003). The estimated rate of exposure per capita due to animal antibiotic use (AAU rate), λ, is also 70% lower than the prior mean assumed in Table 1. The colonization rate θ falls by 30%, but the recolonization rate φ increases by 30% compared to their values in Table 1. These changes indicate that, when the parameter estimates are treated as uncertain rather than as deterministic inputs, Bayesian inference indicates that some of the prior estimates of their values in Table 1 need to be revised to increase consistency with historical data.

Table 2: Prior and Posterior Means of Parameter Values

Parameter	**Smith Prior**	**Posterior (conditioned on low historical resistance rates) (d = 1.5%)**
λ	0.001	0.0003
μ	10^{-6}	$1.033*10^{-6}$
α	0.1	0.1099
θ	0.001	0.0007
σ	0.003	0.0031
γ	0.007	0.0073
φ	0.003	0.0039
ρ	0.003	0.0008
η	10^{-5}	$1.057*10^{-5}$
β	0.5	0.4664

4.2 Revised Prescription Rate for QD

The results in Table 2 suggest a need to re-examine the detailed assumptions behind the value of ρ, the daily prescription rate, in Table 1. It is an estimate for vancomycin, which is far more frequently prescribed than QD. To apply the model in Figure 1 to QD-resistant VREF, a lower value appropriate for the daily prescription rate for QD must be used instead. A plausible correction, confirmed for consistency with the data by repeating the above analysis, was made by assuming that the daily prescription rates of the two drugs vancomycin and QD in the patient population are proportional to their treatment days per person-year in the US. While vancomycin use in the US is estimated as approximately 0.04 treatment-days of therapy per person-year (Pharmacia & Upjohn, 2000), 2001 sales information from Aventis (distributor of Synercid™ at that time) equates to at most 96,667 treatment-days (assuming that retail prescription costs for Synercid approximate $300/day (Aventis, 2002). The US population in mid-2001 was 284,796,887, providing at most an average of 0.00034 Synercid™ treatment-days of therapy per person-year. The ratio of QD to vancomycin use is thus estimated as approximately 0.00034/0.04 = 0.0085, which we round up to 0.01. Thus, the estimated prior mean prescription rate of $\rho = 0.003$/day for vancomycin in Table 1 is revised downward to $\rho = 0.00003$/day for QD.

The initial round of Bayesian estimates thus succeeded in identifying that the value initially assumed for ρ needed to be reduced, although the final revision was based on other available statistical data rather than on Bayesian updating alone. Subsequent rounds of estimation were performed starting from the corrected value.

4.3 Human Health Risks vs. Key Parameter Values

Table 3 shows results for the revised mean value of $\rho = 0.00003$ just described, using a series of different values for the detection threshold, *d*. Significant changes from prior to posterior values now occur only for the exposure rate due to AAU, λ, which is reduced to less than half of its value in Table 1. Table 4 shows sensitivity analysis results for values of the prescription rate, ρ, ranging from 3×10^{-5} (estimated Synercid™ use rate for the United States in 2001) to 4×10^{-3} (slightly above the Smith *et al.* parameter estimate for ρ in Table 1) for two values of λ, the transfer rate due to animal antibiotic use: 0.001 (Smith *et al.* estimate) and 0.0005 (posterior mean).

Table 3: Posterior Values for Varying Prevalence

Parameter	**Prior**	**d=1.5 (n = 4830)**	**d=2 (n = 5604)**	**d=2.5 (n = 6420)**
λ	0.001	0.0004	0.0005	0.0005
μ	10^{-6}	0.0978×10^{-6}	0.994×10^{-6}	1.01×10^{-6}
α	0.1	0.1049	0.1037	0.1039
θ	0.001	0.0008	0.0008	0.0008
σ	0.003	0.003	0.003	0.003
γ	0.007	0.007	0.007	0.007
φ	0.003	0.003	0.003	0.003
ρ	3×10^{-5}	2.945×10^{-5}	2.944×10^{-5}	2.956×10^{-5}
η	10^{-5}	0.99×10^{-5}	0.994×10^{-5}	1.013×10^{-5}
β	0.5	0.4993	0.4980	0.5015

n = number of simulations out of 10,000 producing parameter values leading to a prevalence less than d%.

Each value in Table 4 represents an average over 100 random simulation runs, and thus the probabilities can be given binomial confidence bands with parameters n = 100 and p = tabulated value. (Since all numbers in Table 4 are based on simulation runs, confidence intervals can be made arbitrarily small by increasing the sample size, i.e., the number of runs, but n = 100 suffices to reveal the patterns in the simulation output, described next.) A second set of experiments varied the population size, N, as shown.

The second line of Table 4 (in bold) corresponds to the input parameter values of Smith *et al.* from Table 1 (except for the revised prescription rate ρ, which is now based on QD data.) The bottom line (also in bold) corresponds to the parameter values based on the data and analyses described above. The most striking result from this Bayesian uncertainty analysis is that *the estimated (posterior mean) value of R_0 falls from 1.98 to*

0.02. Using the QD-specific value of ρ reduces the probability of a rapid, sustained increase in SREF among human patients, leading to a positive sustained equilibrium endemic level, from essentially 1 to essentially zero. It changes the model-predicted quantitative answer to Shnayerson and Plotkins' (2002) important rhetorical question, "How long would it be before Synercid-resistant VRE was ubiquitous? Before the new miracle drug was dead?" from "Less than 5 years" to "It will never happen".

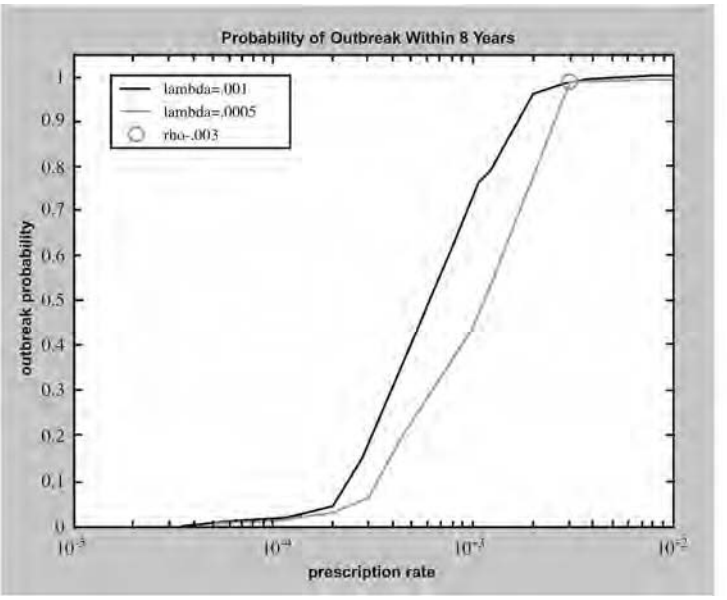
Table 4. DES Model Results for SREFs

N	λ	ρ	R_0	Pr(Outbreak in < 8 years)	Mean out-break size	Avg. SREF cases per patient-year
100	0.001	0.004	2.63	1	14.4	0.478
100	**0.001**	**0.003**	**1.98**	**0.99**	**8.3**	**0.2416**
100	0.001	0.002	1.34	0.96	3.4	0.0676
100	0.001	0.001	0.66	0.73	1.4	0.0094
100	0.001	0.0003	0.204	0.16	1.1	0.0018
100	0.001	0.0002	0.13	0.05	1.0	0.0011
100	0.001	0.0001	0.07	0	1	0.0006
100	0.001	0.00003	0.02	0	1	0.0001
100	0.0005	0.004	2.63	0.99	13.8	0.3819
100	0.0005	0.003	2.0	0.99	7.6	0.1795
100	0.0005	0.002	1.3	0.79	3.0	0.0373
100	0.0005	0.001	0.66	0.43	1.4	0.0052
100	0.0005	0.0003	0.20	0.06	1.1	0.0007
100	0.0005	0.0002	0.13	0.03	1.1	0.0005
100	0.0005	0.0001	0.07	0.01	1.1	0.0004
100	0.0005	0.00003	0.02	0	1	0.0001
1000	0.0005	0.00003	0.02	0.02	1.0	0.00008
700	0.0005	0.00003	0.02	0	1	0.00008
500	0.0005	0.00003	0.02	0	1	0.00009
250	0.0005	0.00003	0.02	0	1	0.00006
50	0.0005	0.00003	0.02	0	1	0.00013
25	0.0005	0.00003	0.02	0	1	0.00010
15	**0.0005**	**0.00003**	**0.02**	**0**	**1**	**0.00009**

Even if continued VM use does not cause a resistance endemic, it might still increase the risk or frequency per year of resistance outbreaks in intensive care units. Defining an outbreak conservatively as more than one amplified case in an ICU at the same time, the probability of outbreak can be estimated for any set of input parameter values as the fraction of simulation runs for those values that produce an outbreak (within the 8-year simulation

time horizon). For the estimated most plausible input parameter values (bottom row of Table 4, $\rho = 0.00003$ and $\lambda = 0.0005$) the probability of an outbreak is 0 (to 3 significant digits, based on simulation sample sizes) except for N = 1000 beds. [Arguably, for such a large N, two simultaneous cases should not necessarily be considered an outbreak, as they are likely to be unrelated. The average ICU size in the US is N = 15 beds (Rapoport, *et al.*, 2003); N = 700 beds is assumed in the Smith *et al.* model.] Figure 2 plots outbreak probability against $\log_{10}(\rho)$. It shows that increasing λ causes the outbreak probability to rise later but more sharply. Extensive simulation shows that outbreak probability is negligible for $\rho \leq 10^{-4}$, which is roughly 1/30 the current vancomycin prescription rate and perhaps 3 times the 2001 prescription rate for Synercid in the US.

Figure 2: Outbreak Probability as a Function of QD Prescription Rate



4.4 Estimated Individual and Population Risks

For individual risks, the average number of cases per patient-year in the rightmost column of Table 4 reflects the annual rate at which an individual patient in an ICU is expected to acquire a case of SREF, computed as the ratio of the number of simulated patients entering the *amplified* state during a simulation run to the number of simulated patient years. Interestingly, the ICU population size N had no significant effect on cases per patient-year. The average "incident size" in Table 4 is defined as the time-averaged number of amplified cases in periods with more than zero

amplified cases. (Following Smith *et al.* (2002, 2003), we only consider SREF infections in ICUs, since QD resistance is a risk only to patients seriously ill with VREF, which occur almost exclusively in the ICU.)

Table 4 shows that the best estimate (bottom row) of the risk of a SREF case per person-year spent in an ICU is about 0.0001, more than 2400 times smaller than the values estimated using the parameter estimates in line 2 (and Table 1). The average stay in an ICU in the US is on the order of 10 days (AHA, 2003), so the individual risk of SREF per visit to the ICU is approximately $(10/365) \times (0.0001) \approx 3 \times 10^{-6}$.

Quantitative mortality impacts of QD treatment failures can be estimated from a case control study by Linden *et al.* (1997), discussed in the Appendix to Chapter 6, which found that 5/20 (25%) of cases given QD had VREF-associated mortality while 17/42 (40.5%) controls receiving alternative treatment had VREF-associated mortality. This suggests an excess mortality rate attributable to not being given QD (and hence for QD resistance) of 15.5% if availability of effective alternative therapies (such as linezolid) is ignored. Multiplying 3×10^{-6} SREF treatment failures per patient per ICU visit by 0.155 as a conservative estimate of the mortality rate of Synercid-resistant VREF patients due to resistance, the individual mortality risk from SREF per hospitalization for the most at-risk group (ICU patients) is unlikely to exceed 4×10^{-7}.

Converting SREF cases per ICU patient-year to cases per capita in the population gives a view of approximate societal risks. The average number of beds in an ICU in the US is approximately 15 (Rapoport, *et al.*, 2003). In 2001 there were a total of 4,880 general hospitals in the US with 849,168 hospital beds (AHA, 2001). If every hospital had an average ICU, a reasonable upper bound on the number of ICU beds in the US would be 15 x 4880 = 73,200. The average number of cases per patient-year for 15-patient ICUs in Table 4 is 0.00009 (see Table 2). Assuming conservatively that ICU beds in the US are always full gives an estimated potential of at most 6.58 cases of SREF per year (95% CI = 0 – 10.98). Multiplying by the above mortality rate if 0.155 for SREF patients due to resistance yields a plausible upper bound estimate of at most 6.58 x 0.155 = 1.02 annual mortalities in the US due to Synercid™ resistance (95% CI = 0 – 1.70).

5. DISCUSSION

This chapter has introduced a version of Bayesian Monte Carlo (BMC) uncertainty analysis that emphasizes acknowledging enough uncertainty in parameter values so that data can identify where initial estimates need to be revised. By avoiding a stronger commitment to the interpretation of prior

distributions for the parameters, we allow a form of data-driven probabilistic sensitivity analysis. We have also shown that BMC can be applied to resolve conflicting intuitions about whether a long history of no emergence of resistance should be considered reassuring in a stochastic dynamic system model of the random emergence of resistance due to selection pressures from animal and human drug use, even though, for some combinations of values of the uncertain input parameters, rapid emergence of resistance in the near future is almost certain. In this particular case, the answer turns out to be that past history *should* be regarded as highly reassuring: within the context of the model, it implies that combinations of parameter values that will lead to future emergence of resistance have essentially zero probability (i.e., they are inconsistent with the data that higher levels of resistance have not already emerged). However, this conclusion is specific to the virginiamycin example and data reviewed. For other antibiotics and bacteria, an analogous BMC analysis might lead to a different conclusion. The possibility of long-delayed emergence of high levels of resistance among patients cannot be denied in general, but the BMC approach provides a constructive approach for assessing or bounding its probability given historical data, despite the complexity (and potentially increasing hazard rate for emergence) of the discrete-event stochastic simulation model.

In retrospect, the full power of the BMC approach is not needed to establish the main practical conclusions from this study: that the prescription rate for QD is much less than for vancomycin and that making this adjustment reduces R_0 well below 1 and reduces the risk of substantial future increases in human resistance levels or resistance outbreaks if VM continues to be used from near 1 to near 0. Deterministic sensitivity analysis can make the same points. The major practical contribution of BMC was to automatically highlight which specific quantities (in this case, just ρ and λ) required revision to be consistent with historical data. This has led to different insights than the deterministic sensitivity analysis of Barber *et al.* (2003). Their analysis, unguided by BMC, focused on other parameters and reached the qualitative conclusion that the equilibrium endemic prevalence level predicted by the model might be much smaller than in the Smith *et al.* analysis, e.g., a few percent instead of 50%. By contrast, the additional quantitative assessment made possible by the BMC approach showed that the probability of a self-sustaining endemic of any positive size is approximately zero.

Finally, the BMC analysis has highlighted the fact that the probability of high levels of human resistance to Synercid™ (QD) emerging in future is very sensitive to human use rates (i.e., the prescription rate ρ) and is comparatively insensitive to plausible changes in exposures from animal antibiotic use (λ) (see Figure 2).

6. SUMMARY AND CONCLUSIONS

How reassuring is the past about the future? This chapter has applied quantitative risk assessment methods to help find out, using human health risks from VM and the nearly identical human antimicrobial quinupristin-dalfopristin (QD) as a case study. A dynamic simulation model was used to predict risks of emerging resistance to human antimicrobials in human populations from given input assumptions. Bayesian Monte Carlo (BMC) uncertainty analysis allowed past data to constrain and inform selection of input parameter values and thus to predict possible future resistance patterns that are consistent with historical data. The results showed that health risks from VM use in food animals are highly sensitive to the human prescription rate of QD. For realistic prescription rates, quantitative risks are less than 1×10^{-6} even for members of the most-threatened (ICU patient) population, while societal risks are less than 1 excess statistical death per year for the whole United States population.

Such quantitative estimates complement assessments that discuss the possibility of future "resistance epidemics" (or endemics) without quantifying their probabilities. For example, Smith *et al.* (2003) warned that

> "Virginiamycin, another streptogramin, threatens the efficacy of QD in medicine because streptogramin resistance in enterococci associated with food animals may be transferred to *E. faecium* in hospitalised patients. …To provide a sound basis for policy, we have reviewed the epidemiology of *E. faecium* and streptogramin resistance and present qualitative results from mathematical models. These models are based on simple assumptions consistent with evidence, and they establish reasonable expectations about the population-genetic and population-dynamic processes underlying the emergence of streptogramin-resistant *E. faecium* (SREF). Using the model, we have identified critical aspects of SREF emergence. We conclude that the emergence of SREF is likely to be the result of an interaction between QD use in medicine and the long-term use of virginiamycin for animal growth promotion. Virginiamycin use has created a credible threat to the efficacy of QD by increasing the mobility and frequency of high-level resistance genes. The potential effects are greatest for intermediate rates of human-to-human transmission (R_0 approximately equal 1)."

This analysis and similar qualitative concerns have prompted many groups to call for prompt regulation and bans of virginiamycin (as well as other animal drugs) to protect the efficacy of QD in human medicine (APUA, 2002; Keep Antibiotics Working, 2002).

The quantitative risk assessment presented here, using BMC uncertainty analysis and revised parameter estimates consistent with

available data specifically for virginiamycin and Synercid™, suggests a different perspective. The risks in question are expected to be less than 1 statistical death per year in the whole US population and not more than about 4×10^{-7} individual risk per ICU hospitalization even for the most at-risk individuals (those in ICUs.) For such small risks, it is unlikely that "interaction between QD use in medicine and the long-term use of virginiamycin for animal growth promotion" has any detectable effects on human health. More importantly, the emergence of higher levels of SREF in the human population in the future due to continued use of virginiamycin in animals and Synercid™ (QD) in people has been shown to have approximately zero probability (according to the model with any of the posterior parameter values in Table 3), and so the risk that continued use of virginiamycin in animals will shorten the useful life of QD as a human drug is also approximately zero in this model. This BMC-derived information complements the qualitative conclusions quoted above by showing that the theoretical possibility that they refer to is very unlikely in light of historical data, at least in the context of the Smith *et al.* model in Figure 1. As Smith *et al.* noted, the model has not been validated, and so conclusions based on it may not be realistic. However, the major new insight contributed by the BMC analysis is that the model gives no reason to suspect that continued use of QD poses a significant threat to human health.

Uncertainty and sensitivity analyses (see Tables 2-4 and Figure 2) show that our findings are driven primarily by the use of a realistic value of the prescription rate for QD. Although the BMC uncertainty analysis suggested other potential revisions in parameter values to reflect the information that high rates of human resistance to QD have not yet emerged (Tables 2-4), the small prescription rate of QD in the United States limits the potential for emergence of resistance and for consequent harm to human health. This finding is robust to a wide range of uncertainties in the values of other parameters. The systems dynamics model in this chapter supports the conclusion from Chapter 6, that potential human health risks from continued use of virginiamycin appear to be relatively small. The following chapter turns to potential human health benefits from continued use.

This page intentionally blank

Chapter 8

Potential Human Health Benefits of Animal Antibiotics

1. INTRODUCTION

Risk management of food-animal antibiotics has reached a crucial juncture for public health officials worldwide. While withdrawals of animal antibiotics previously used to control animal bacterial illnesses are being encouraged in many countries, the human health impacts of such withdrawals are only starting to be understood. As discussed in Chapter 1, increases in animal and human bacterial illness rates and antibiotic resistance levels in humans in Europe despite bans on animal antibiotics there, coupled with declining illness rates in the United States despite continued use of animal antibiotics, suggest a need to carefully examine how strongly and in what ways animal antibiotic uses affect human health.

This chapter continues the quantitative investigation of potential human health impacts of animal antibiotic uses begun in Chapter 6. It extends the attribution-based, product-of-factors approach to help assess the potential human health benefits, as well as the human health risks, of continued use of animal antibiotics. Virginiamycin (VM) is again used as the main case study for developing and illustrating a practical analytic framework for assessing the potential human health impacts – both good and bad – of animal antibiotic use. The framework is also applied, albeit more briefly, to macrolides and fluoroquinolones, which have also been recommended for withdrawal from use in food animals in several countries. In contrast to the treatment in Chapter 6, this chapter assesses impacts of interventions on *total* human illness rates – specifically including illnesses caused by susceptible bacteria as well as those caused by resistant bacteria.

A major conclusion is that, if banning or restricting use of common animal antibiotics causes more pathogens to reach consumers, then the resulting human health risks might well far exceed the potential human health benefits due to decreased resistance. While lack of hard data on how withdrawing animal antibiotics affects average human illnesses per serving of meat precludes a deterministic conclusion, the partial information available now indicates that bans on animal antibiotics could cause thousands of times more human illness-days per year than they would prevent. This conclusion, though tentative, is robust to several important scientific and modeling uncertainties.

The validity of this conclusion depends on two main factors, both of which are uncertain: (1) The extent to which ceasing an animal antibiotic use will increase microbial loads of pathogens in servings; and (2) The dose-response relation between increased microbial loads in servings and resulting increases in human illnesses. Faced by these important uncertainties, we apply the bounding structural equation modeling approach of Chapter 5 to obtain rough bounds on the potential human health risks and benefits of banning continued animal antibiotic use. These bounds are developed contingent on the simplifying assumption that human health risks are directly proportional to average microbial loads ingested, and sensitivity analysis is used to show how the results change for a non-linear dose-response model.

Because the potential risks of bans appear to be so much larger than the potential benefits, there is a high value of information (VoI) from collecting additional data before taking action. Specifically, additional studies that compare microbial loads in carcasses and/or servings from antibiotic-treated and antibiotic-deprived animals (similar to the study of Russell, 2003 for *Campylobacter* in processed chicken carcasses) would be most valuable in providing the critical factual information needed to determine whether animal antibiotic bans are more likely to help or to harm human health.

The following sections first develop and illustrate an attribution-based Rapid Risk Rating Technique (RRRT) approach to human health impacts assessment for virginiamycin, then apply it to quantify bounds on human health risks and benefits for macrolides and fluoroquinolones.

1.1 A Risk Management Dilemma for Virginiamycin

Virginiamycin is one of several antibiotics (including avoparcin and narasin) that are effective in controlling bacterial enteritis in food animals and promoting uniformity in weight of animals at slaughter (e.g., George, 1982). Veterinary and agricultural experience suggest that more uniform weights and decreases in bacterial illnesses that cause under-weight chickens at slaughter are associated with lower loads of microbial pathogens such as

Campylobacter jejuni and *Salmonella* on processed carcasses (Dawe, 2004). Thus, use of animal antibiotics for prophylaxis and growth promotion may help to reduce *C. jejuni* and other pathogens (including *Salmonella*) in processed retail meat. If so, then withdrawing animal antibiotics might increase rates of animal bacterial illness and human foodborne illnesses.

Empirically, key animal and human zoonotic bacterial illness rates and antibiotic resistance levels in humans increased in Europe immediately following bans of animal antibiotics (including VM) used as prophylactics and growth promoters (see Chapter 1, Figure 1 and *Eurosurveillance,* 2002), as well as during earlier periods of voluntary restrictions. At the same time, campylobacteriosis rates declined dramatically in the United States, which continued to use animal antibiotics (CDC 2000, 2003; Stern and Robach, 2003). Such data have led some to question long-standing science policy assumptions and assertions that withdrawing animal antibiotics will promote human health, or, conversely, that continued use of animal antibiotics increases human antibiotic resistance or illness rates (e.g., Cromwell, 2002; Casewell *et al.*, 2003; Phillips *et al.*, 2004). Yet, many scientists involved in making policy recommendations claim that animal antibiotic bans logically should be, and in fact have been, successful, at least as measured by reductions in resistance in food animals and in healthy humans (e.g., Wegener *et al.*, 1999, 2003). Even when withdrawing animal antibiotics has been followed by deteriorations in human health and increases in antibiotic resistance among human patients, it has also been followed by reductions in resistance to antibiotics among harmless bacteria, and this has sometimes been proposed as a measure of the success of the bans. Also, it is not yet clear *why* human illness and resistance rates have increased Europe. Changes in chicken processing, preparation, and consumption patterns (e.g., substitution of fresh for frozen chicken or changes in imports) have been conjectured as possible explanations, but no thorough quantitative analysis has yet shown what role, if any, the animal antibiotic bans have played.

This mixed evidence to date creates a dilemma for public health officials and regulators. Which creates a larger net public health benefit: withdrawing animal antibiotics to reduce selection pressure for resistance in bacteria, or continuing their use to reduce the incidence of foodborne illness and consequent need to treat some human patients with human antibiotics? The answer is not intuitively obvious. Most previous risk assessments have focused only on the risks from resistance without comparing them to the benefits from prevention of illnesses caused by susceptible bacteria (e.g., FDA-CVM, 2001, 2004). This chapter seeks to develop answers that are clear, robust and credible enough to be useful to decision-makers while taking realistic complexities and uncertainties into account.

2. AN RRRT FRAMEWORK FOR ASSESSING IMPACTS

This section describes a Rapid Risk Rating Technique (RRRT) for estimating quantitative impacts of animal antibiotic uses on annual rates of adverse human health effects in a population exposed to bacteria via the food chain and perhaps other paths. These impacts are human health *risks* if they increase the rates of adverse health effects and *benefits* if they reduce them. Both are expressed in units of change in expected numbers of adverse consequences of different severities per capita-year (for individual risks) or per year (for population risks) caused by the risk management option(s) being evaluated. Severities of outcomes may be indicated by severity classes (e.g., mild, moderate, severe, or fatal, as defined by Buzby, *et al.*, 1996) or, if desired, by quality-adjusted life-years (QALYs) if the required assumptions are acceptable (Hazen, 2003). As in Chapter 6, we will focus on risks from highly (vanA) vancomycin-resistant *E. faecium* ($VREF_A$) infections in ICU patients as the population primarily affected by QD-resistant VREF infections. Health consequences considered include severe illness only, severe illness with treatment failure but not mortality, and fatal cases of $VREF_A$. (Patients without serious illnesses are not normally at risk of $VREF_A$ infections.) For fluoroquinolones, macrolides, and virginiamycin, risks of campylobacteriosis in the general population will also be considered.

The basic logic of the RRRT approach for health risk and benefit assessment is to compare the expected incremental numbers of adverse human health consequences per year (a) *caused by* an animal antibiotic use (due to increased selection of resistance determinants and/or resistant bacteria); and (b) *prevented by* the animal antibiotic use (due to reductions in animal illnesses and resulting reductions in microbial loads reaching consumers via meat products). Use of expected number of events per year to quantify risk is justified for sporadic illnesses that occur independently or with only weak statistical dependence in large populations under the conditions of Poisson or compound Poisson approximations; see Chapter 2. For commensals, the top-level RRRT formulas are as follows:

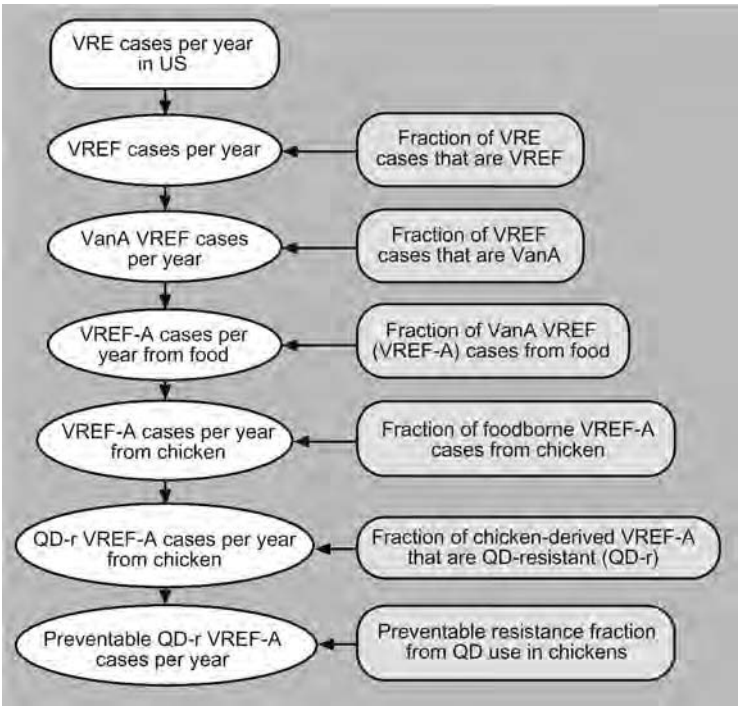
- RISK from continued animal drug use = (preventable resistant illness cases caused per year) × (adverse clinical consequences, such as incremental illness-days, per resistant case)
- BENEFIT from continued animal drug use = (prevented illness cases per year) × (adverse clinical consequences avoided per case prevented)
- NET HEALTH IMPACT of continued animal drug use = BENEFIT – RISK = human health harm prevented – preventable human health harm caused by continued use. ("Preventable" here means preventable by

discontinuing the animal drug use. Analogous definitions hold for risk, benefit, and net health impact of introducing a new animal drug use.)

In these formulas, all quantities denote expected values. The formulas for human health RISK and BENEFIT each has the form: (expected cases) × (expected consequence per case). Justification for using these products of expected values to obtain the expected total harm caused or prevented by continued use, respectively, follows from general results for sums of random numbers of random variables (representing a random number of illnesses, each incurring a random number of quality-adjusted life-years (QALYs) lost, illness-days incurred, etc.) (Feller, 1968). Adverse clinical consequence terms may be expressed as vectors of expected illness-days in each severity class per illness case; expected number of fatalities per illness case, etc.; or, for an aggregate summary, as average QALYs lost per illness case.

To estimate the RISK and BENEFIT formulas from data, each term is decomposed into a product of more-easily calculated factors. Figure 1 shows the structure of such a calculation for the quantity "preventable resistant illness cases per year" appearing in the formula for RISK.

Figure 1: Calculation Logic for Preventable Resistant Illness Cases Per Year



The value of each random variable, represented by an elliptical node in this influence diagram, is calculated as the product of the values of the nodes that point into it. The product form entails no loss of generality, as *any* joint probability density of values of variables, say, Pr(a, b, c, …, y, z), can be factored as a product in the form: $\Pr(a) \times \Pr(b \mid a) \times \Pr(c \mid a, b) \times \ldots \times \Pr(z \mid a, b, c, \ldots y)$. In particular, as noted in Chapter 6, a conjunction of conditions (e.g., that an infection be $VREF_A$, from chicken, QD-resistant, etc.) can be expressed as such a product. The sequence of inputs on the right side of Figure 1 is such a chain of conditional relations, with each factor being conditioned on all those that precede it. (Of course, if some factors are irrelevant, then they can be omitted from the conditioning.) Values of the input nodes (i.e., nodes having only outward-pointing arrows) are estimated from available data, as detailed below.

As discussed in Chapter 5, uncertainties in the inputs can be propagated through this diagram using Monte Carlo uncertainty analysis, e.g., using the Analytica™ influence diagram software used to draw it (http://www.lumina.com/). A simpler approach is to use upper-bound estimates for uncertain quantities. The multiplicative architecture of the model implies that evaluating *any* subset of the input nodes and multiplying them by the first (VRE cases per year in the United States) will give an upper-bound estimate of the final output quantity. Multiplying by further fractions between 0 and 1 can only reduce the estimate further.

The other terms required to calculate the BENEFIT and RISK formulas can be expanded similarly into influence diagrams representing products of factors that are ultimately estimated from data (or sums of such products.) Table 1 in the next section summarizes the parameters and estimated values used in calculating RISK. Table 2 summarizes the parameters used in calculating BENEFIT. Briefly, if a ban causes an increase ΔF in the fraction of chicken servings from "ill" or "high-risk" (e.g., necrotic-enteritis positive (NE^+)) flocks instead of healthy or low-risk (NE^-) flocks, and if each such serving creates an incremental probability ($P^+ - P^-$) of causing illness (campylobacteriosis), with an average health impact per illness of Q illness-days or QALYs, then the expected human health impact caused by preventing the increase ΔF in animal illness prevalence is:

BENEFIT from continued use of a drug use that prevents ΔF incremental fraction of high-risk servings = $\Delta F \times (P^+ - P^-) \times$ MNQ illness-days

where N = average chicken servings per capita-year and M = number of people in the population.

To a public health risk manager, the main question of practical interest is the sign of (BENEFIT – RISK), i.e., is the net human health impact from

continued antibiotic use in animals positive or negative compared to the results of a ban? After addressing this question for the particular case of VM use in chickens, the results will be used as a point of departure for considering how other bacteria, food animals, pathways, and health effects might change the answers. This approach allows rapid consideration of many other factors that may affect (BENEFIT – RISK) but without greatly changing the key decision-relevant findings.

3. DATA FOR RRRT RISK-BENEFIT CALCULATIONS

Table 1 summarizes the data sources and calculations of the RISK component using the RRRT framework. Chapter 6 discussed the derivations of the first six parameter estimates in the "Values for USA" column in Table 1. Several conservative estimates (e.g., fractions set equal to 1, maximizing estimated risk) are used where data are missing or inadequate. The rationale is that confidence in the conclusion that the sign of (BENEFIT – RISK) is positive is strengthened if we deliberately bias the analysis against it by choosing estimates of uncertain quantities that tend to over-estimate RISK and underestimate BENEFIT. If the result is still that BENEFIT > RISK, as in the following analysis, then using more realistic (less biased) estimates of the uncertain quantities when and if uncertainties are reduced will tend to further strengthen this conclusion.

Table 2 summarizes the calculation of the human health BENEFIT from continued VM use. Key components of the calculation are as follows.

3.1 Calculation of ΔF

Several animal antibiotics, including macrolides and streptogramins (VM), are effective against various bacterial illnesses in chickens, including necrotic enteritis (NE) caused by *Clostridium perfringens* (George *et al.*, 1982; Brennan *et al.*, 2001). Withdrawing these animal antibiotics nationwide may therefore increase the fraction of servings from ill flocks and birds, e.g., NE^+ flocks or flocks with other bacterial illnesses having similar impacts on increasing pathogen loads on processed carcasses (e.g., Dawe, 2004). ΔF denotes the size of this fractional change. [By contrast, more limited withdrawals, such as in only one or a few farms, can have little or no effect on increasing disease rates (Engster, *et al.*, 2002) if continued use elsewhere holds infectious disease levels in check. Infectious disease dynamics typically require a critical threshold, R_0, of infected animals to be passed before epidemics can spread widely.]

Table 1: Summary of Risk Assessment Calculations for VM

PREVENTABLE QD-RESISTANT VREF CASES PER YEAR		
Factor	**Values for USA**	**Data Sources**
Expected number of VRE cases/year in ICU population	37,483 (From Chapter 6, Sect. 2.1)	NNIS, 2001; Lawton et al, 2000; AHA, 2001
Fraction of VRE cases that are VREF	0.71, 0.78, 0.95, Median: **0.78**	SNJ, 2000; Clark *et al.*, 1993
Fraction of VREF cases with vanA (high level) resistance, $VREF_A$	0.73, 0.79, 0.83 Median: **0.79**	Eliopoulos1998; Clark., 1993; Jones, 1995
Fraction of $VREF_A$ cases in hospitals that could have originated from food	< 0.17 = Proportion not nosocomial	Austin *et al.*, 1999; Thal *et al.*, 1998
Fraction of $VREF_A$ cases from food that might have come from chickens	0 to 0.12 based on genogroups	Willems *et al.*, 2000 Willems *et al.*, 2001
Fraction of foodborne $VREF_A$ cases that are QD-resistant (QD-r $VREF_A$)	0.011 (Chapter 6)	Eliopoulos *et al.*, 1998, Jones *et al.*, 1999
Fraction of foodborne QD-r $VREF_A$ cases with QD-resistance caused by VM use in chickens	< 1	Upper bound
Preventable resistance fraction = fraction of foodborne QD-r $VREF_A$ (i.e., $SREF_A$) cases that might be prevented if VM use in animals ceased	≤ 1	Upper bound. (Cox and Popken, 2004 estimate 0.68 within 5 years)
CLINICAL CONSEQUENCES PER QD-RESISTANT VREF CASE		
Fraction of QD-r $VREF_A$ cases not treated successfully with linezolid or with other non-QD antibiotics	≤ 0.074 (0.074 = failure fraction for linezolid alone)	Linden *et al.*, 2002
Fraction of $SREF_A$ cases not treated successfully with non-QD antibiotics that are then prescribed QD and that fail to respond normally	< 1	Upper bound
Fraction of $SREF_A$ cases prescribed QD that fail to respond normally *because of* the QD-resistance	0.7	Linden *et al.*, 2002 Moellering *et al.*, 1999
Increased mortality probability due to QD resistance	0-0.11 (Chapter 6, Appendix)	Linden *et al.*, 1997
QALYs lost per non-fatal case	0.04 QALYs, 14.6 illness-days (*ibid*)	Webb *et al.* 2001
Average QALYs lost per fatal case	21.7 (Appendix)	Webb *et al.* 2001
Preventable excess mortalities per year < 0.03 = 37483*0.78*0.79*0.12 *0.011*0.17*0.074*0.7*0.11	< **0.03 mortalities / yr**; 0.03/0.11 = **0.27 cases/yr**	Product of above lines (using upper bounds)
Preventable excess morbidity QALYs per year = 37483*.78*.79 * .12*.011*.17*.074*.7*.89*0.04	**0.001** QALYs (corresponds to **3.5** illness-days/year)	Cases per year * 0.89 non-fatal*0.04 QALYs each
Preventable QALYs lost per year = (0.03*21.7 QALYs/mortality) + 0.001 from morbidities	< **0.65 QALYs**	

To quantify ΔF, we note that NE rates in several countries increased sharply, if transiently, following bans on VM, macrolides, and other animal antibiotics used as prophylactics and growth promoters, before settling to new levels with increased use of therapeutic drugs (e.g., Lovland and Kaldhusdal, 2001; Madsen and Pederson, 2000; Veterinary Laboratories Agency, 2004). For Denmark, Madsen and Pederson (2000) reported that: "In 1998, necrotic enteritis was diagnosed in 25 out of 1,700 Danpo flocks as compared to a few flocks annually before the discontinued use of antibiotic growth promoters." Thus, the incidence of NE in Denmark may have increased by about 23/1700 or 1.35 percentage points from 1997 to 1998. However, macrolides (not used much in Denmark), avilamycin (AV) and Zinc Bacitracin have similar effects to VM, and VM accounted for only 0.32 of these drugs by weight in 1996 (DANMAP, 1997). If a kg of each product is similarly effective in preventing NE, the increase in NE incidence attributable to withdrawing VM would be about $\Delta F = 0.32 \times 1.35 = 0.43$ percentage points.

Table 2: Parameters For RRRT Baseline Benefit Assessment For VM

Symbol	Meaning	Baseline value and source
ΔF	Fractional change in prevalence of chicken servings from ill or high-risk [e.g., NE-positive (NE^+)] flocks if current VM use ceases	0.5% = assumed baseline
P^-	Average probability of illness per serving from animals without disease. Includes indirect effects of cross-contamination of other foods. This probability is an average for the whole population; individual risks may vary.	1.3E-5 = (total *C. jejuni* illnesses per year) × (fraction caused by chicken)/(total chicken servings per year)
$P^+ - P^- = (1 + R)*P^-$	Excess probability of illness per serving from NE+ flocks. (Includes cross-contamination effects)	1.2E-4 (for linear no-threshold dose-response model, microbial load ratio ≈ 10, from Russell, 2003)
M	Average number of servings of food commodity ingested per capita-year	38 FDA-CVM, 2001, Cox and Popken, 2002 for fresh chicken
N	Number of people in population	292E6 (U S Census)
Q	Average human health harm (e.g., days of illness or QALYs lost) per case. Interpreted as "severity" of a case.	6.13 days (Marano *et al.*, 2000); 0.0043 QALYs, ≥ 8E-5 fatalities per case (Buzby, *et al.*, 1996)
Risk created by ban	**41,016** = (0.005*1.2E-4 * 38*292E6 * 6.13 excess illness-days) per year = 6691 additional cases*6.13 days/case.	$[\Delta F(P^+ - P^-)]MNQ$ = 0.53 fatalities, 28 QALYs lost, 41,016 illness-days per year

In the United States, where avilamycin is not used, the role of VM would presumably be greater; we round it up to **ΔF = 0.5%**. This is the baseline value used in Table 2. However, the true value of ΔF for the United States is quite uncertain. For example, in the United Kingdom., the NE^+ fraction of flocks increased from 0% before the ban to over 20% in the years following it before finally being brought under control by other antibiotics and countermeasures (Veterinary Laboratories Agency, 2004). To account for this uncertainty, we interpret the results in Table 2 as the human health benefit created (i.e., human health harm prevented) *per half-percentage point* of NE^+ flocks (or similarly ill flocks) prevented.

The most general interpretation of ΔF is that it is the fractional change in servings with exceptionally high microbial loads of pathogens such as *Salmonella*, *Campylobacter*, and *E. coli* (i.e., "high risk" servings) caused by a ban on current antibiotic use. In reality, this fraction may change over time in the several years following a ban. Experience in Scandinavian countries and elsewhere suggests that changes in practice and increased use of other antibiotics (such as ionophores) can control necrotic enteritis in chickens, perhaps after transient increases, even after a withdrawal of virginiamycin and other antibiotics used as growth promoters. How well these results extend to other countries that have different climates and conditions and higher rates of poultry diseases (NE or others, such as airsacculitis) remains unknown. Thus, it is useful to interpret ΔF as the fractional change in high-risk servings, having elevated pathogen loads creating increased risk for human illness, even though the details of the changes in microbial ecology caused by a ban and, in turn, causing ΔF may differ among countries and over time within a single country.

For concreteness, the following calculations focus specifically on increased loads of *Campylobacter* in high-risk servings as compared to regular servings. However, similar logic applies to other pathogens as well, as discussed briefly following the *Campylobacter* calculations.

3.2 Calculation of P^-

P^- denotes the average risk of a campylobacteriosis illness case per chicken serving from a healthy (e.g., NE^-) flock, including possible effects of cross-contamination of other foods. Since nearly all chicken-borne *C. jejuni* cases currently come from healthy (NE^-) flocks, P^- can be estimated as follows:

P^- = (total chicken-caused cases)/(total servings) = [(13.4E-5 reported campylobacteriosis cases per capita-year, from CDC 2003) × (38 assumed cases per reported case, from Mead *et al.*, 1999) × (292E6 people in US, from US Census) × (10% estimated fraction of cases from chickens,

discussed below)]/[(292E6 people in US) × (38.0 servings per capita-year of "fresh" chicken, from Cox and Popken, 2002)] = **1.34E-5** expected campylobacteriosis cases per serving.

Multiplying by the denominator, this corresponds to about (1.34E-05 cases/serving) × (38 servings per capita-year) × (292E6 people in the US) = **1.487E5** estimated chicken-borne cases per year.

The fraction of total cases caused by eating chicken was estimated as about 57% (FDA-CVM 2001) based on pre-1985 data, but *Campylobacter* counts on processed broiler carcasses have since been reduced by perhaps 90% or more (Stern and Robach, 2003). The true fraction of campylobacteriosis cases caused by eating chicken may have fallen from a pre-1995 value of at most 100% to a current value of at most 10%, assuming a proportional reduction in human risk of chicken-borne campylobacteriosis. The baseline calculations in Table 2 assume a value of **0.10**. Section 4 of Chapter 4 provides additional discussion of this value. If a different value is assumed, then the estimated value of P^- will change in proportion.

Calculating ($P^+ - P^-$): Excess Risk per Serving From Ill Chickens

Direct studies of the human health risks of consuming chicken servings from NE^+ flocks as compared to NE^- flocks are lacking for the United States, where almost all flocks are currently NE^-. However, limited data on another poultry disease, airsacculitis (AS), that may have somewhat similar effects (Dawe, 2004; Lovland and Kaldhusdal, 2001) indicate that airsacculitis-positive (AS^+) flocks are associated with increased *Campylobacter, E. coli,* and *Salmonella* on processed carcasses. The increased pathogen loads are caused primarily by greater variability in carcass sizes (see also Engster, *et al.*, 2002) and weakened digestive tracts leading to increased fecal contamination and microbial loads on processed carcasses (Russell, 2003). The mean $\log_{10}$ microbial load of *Campylobacter* colony-forming units (CFUs) before the inside/outside bird wash step of chicken processing for AS^- flocks was 1.09 while the mean for AS^+ flocks was 2.09; thus, the microbial load is about one log (10-fold) higher for the AS^+ flocks, although there is considerable flock-to-flock variability (with increases in only 3 of 5 replicates), so that arithmetic means do not fully represent the change in the microbial load distribution. If a linear no-threshold model is used (i.e., if human campylobacteriosis risk is assumed to be proportional to CFUs per processed carcass) and if the risk of chicken-borne campylobacteriosis is about 10 times greater for carcasses from "high risk" birds compared to regular ones (using the microbial load data of Russell, 2003 as a rough guide) then ($P^+ - P^-$) = ($10 \times P^- - P^-$) = $9 \times P^-$ = 9 × 1.34E-5 = **1.2E-4** is the excess

individual risk of campylobacteriosis per serving from a high-risk (e.g., NE^+ or AS^+) bird.

Sensitivity analyses using alternative dose-response models show that the estimated excess risk per serving from "high-risk" (ill or underweight) birds is sensitive to the use of potentially more realistic non-linear dose-response relations. For example, the FDA Center for Veterinary Medicine suggested a log-exponential model to account for variability in chicken-borne exposures to *Campylobacter* (FDA CVM, 2001). As shown in the Appendix to this chapter, using this exposure-response model produces a revised estimate of: $(P^+ - P^-) = 138.3*P^- = 0.002$, as compared to the above linear-model estimate of 1.2E-4 = 0.00012. Thus, the lower bound on excess risk per serving based on the linear no-threshold model under-estimates the true excess risk per serving by a factor of 0.002/1.2E-4 = 16.7 if the log-exponential model is correct. Using the linear no-threshold model's estimate of 1.2E-4 is consistent with the strategy of choosing estimates that tend to under-estimate the risks from a ban (or the benefits of continued use).

4. RESULTS FOR VIRGINIAMYCIN IMPACTS

Tables 1 and 2 are intended to provide reusable templates, populated with plausible parameter values based on currently available data, for estimating and comparing the values of the human health BENEFIT created (i.e., lost QALYs prevented per year) and RISK caused (i.e., lost QALYS caused per year) by continued use of VM in chicken flocks. For a fractional change $\Delta F = 0.5\%$ of chicken flocks changing from being healthy (e.g., NE^-) to being ill (e.g., NE^+) following a withdrawal of VM, the resulting percentage increase in human campylobacteriosis risks from eating chicken (assuming a directly proportional relation between microbial load of *Campylobacter* at processing and risk of human campylobacteriosis illness) is estimated to be: $(99.5\%) \times 1 + (0.5\%) \times 10 = 104.5\%$ (i.e., 99.5% of chickens would be unaffected and 0.5% would be about 10 times riskier than at present.) This corresponds to an estimated *4.5% increase in chicken-borne campylobacteriosis cases per 0.5% increase in flock illness rates if VM use is withdrawn.* (As indicated above, allowing for log-exponential interindividual variability in infectious doses received from a given microbial load at processing, e.g., due to differences in handling, cooking, and susceptibility, would increase this estimate approximately 17-fold.) If the baseline risk is 1.487E6 estimated cases per year in the United States, with 0.10 being caused by chicken-borne Campylobacter, as estimated above, then the estimated increase in cases per year from withdrawal of VM would be: $(1.487E6) \times (0.10) \times (0.045) = 6691$ additional campylobacteriosis cases per

year. Of these, only a fraction of about 0.006 are expected to be severe (Buzby, *et al.*, 1996), giving an estimate of about 6691 × 0.006 = 40 severe campylobacteriosis illness per year. Buzby *et al.* (Table 2, p. 4) also estimate a fatality rate of at least 200 deaths per 2.5 million cases (and possibly 730 deaths per 2.5 million cases), giving an estimate of at least 6691 × (200/2.5E6) = 0.54 excess deaths.

The baseline calculations in Tables 1 and 2 indicate that withdrawing VM from use in chickens in the United States would prevent not more than 0.65 QALYs lost per year (from less than 0.3 preventable resistance cases, assuming that, in future, QD will still be prescribed in all cases of linezolid failure. This excess case rate corresponds to 0.03 excess fatalities and 3.5 excess illness-days per year). It would be expected to cause over 40,000 excess illness days per year from campylobacteriosis (corresponding to about 6,691 excess cases of campylobacteriosis, 40 of them severe; 0.54 excess deaths; and 28 QALYs lost to illness, based on 0.0043 QALYs per case (Buzby, *et al.*, 1996)) for each half-percent increase in NE-positive (or similarly ill or underweight) chicken flocks. Thus, the expected net human health impact of withdrawing current QD use under these assumptions would be negative: QALYs caused exceed QALYs prevented by over 40-fold, while the fatality ratio is at least 0.54/0.03 = 18; and illness-days caused exceed illness-days prevented by over 40,000/3.5 > 10,000-fold (Tables 1 and 2). From this perspective, the current information and assumptions incorporated into the calculations in Tables 1 and 2 would not justify banning VM use in chickens, but rather suggest that continued use may protect human health.

5. UNCERTAINTY AND SENSITIVITY ANALYSES

To reverse the conclusion that a VM withdrawal would create more cases of campylobacteriosis per year (baseline estimate = 6691) than the number of QD-resistant VREF cases it would prevent (baseline estimate = 0.27), one might seek to increase the estimated fractions in Table 1. For example, suppose that it were assumed that *all* $VREF_A$ infections in hospitals come from VM use in chickens (rather than the baseline estimated fraction of 0.17 × 0.12 = 0.02 in Table 1, based on assumptions that nosocomial cases would not be significantly affected by VM use in chickens and that only human cases with genetic types found in chickens could have come from eating chickens). Then the estimate of preventable QD-resistant $VREF_A$ cases would increase from its baseline value of 0.27 per year to a revised value of 0.27/(0.17*0.12) = 13.2 cases per year. If, in addition, linezolid and other alternatives to Synercid™ were to be withdrawn from the market, or if complete resistance to them emerged, then the cases per year could increase

further, to (13.2)/(0.074) = 178.4. Finally, if the fraction of chicken-derived $VREF_A$ cases that have QD-resistance were also increased by an order of magnitude, from 0.011 to 0.11, then the new estimated number of cases per year, 1784, would be much closer to the estimated prevented campylobacteriosis cases per year, 6691. (Differences in QALYs per case between these two illnesses reduce the differences in their public health impacts, as quantified above; thus, the order-of-magnitude comparisons of cases per year are only a rough guide.)

More generally, the calculations in Table 1 are organized as the product of a base number (37,483 VRE cases per year) multiplied by several fractions that are all between 0 and 1. Increasing any of these fractions (or any two of them, or even any three of them) to their logically maximum possible values of 1 would not increase the baseline estimate of preventable QD-resistant $VREF_A$ cases per year above the estimated preventable campylobacteriosis cases per year, 6691. Thus, despite the uncertainties in the analysis, it appears that this major comparative conclusion is robust to uncertainties or changes in any single assumption, or any small (< 4) set of assumptions, in Table 1.

By contrast, in Table 2, it is only necessary to change ΔF to 0 or $P^+ - P^-$ to 0 to reduce estimated benefits to 0. We have attempted to choose conservative values of these quantities: $\Delta F = 0.005$ instead of several percentage points; and $P^+/P^- = 10$ instead of $P^+/P^- \approx 140$, as estimated for the FDA, 2001 log-exponential model (see the Appendix, last line), as well as assuming that only 0.10 of campylobacteriosis cases are caused by chicken, instead of 0.57 as in FDA, 2001. However, the true benefits could also be as small as zero if $(\Delta F) \times (P^+ - P^-) = 0$, depending on how animal illness rates and microbial loads would change following a VM withdrawal. If NE were to increase sharply, as in Norway following the ban on QD and other growth promoters, then predicted human health harm would increase proportionally to ΔF and the benefits (avoided human health harm) estimated in Table 2 for an assumed ΔF of 0.005 increase in NE^+ flocks might be too small. Thus, the human health risk and benefits estimates in Tables 1 and 2 should be viewed as conservative but uncertain estimates (intended to be probably too high for risks and too low for benefits, to reduce a decision-relevant difference that is already large) that may change as more scientific information about the microbial load and human risk impacts of VM withdrawal become available. While the baseline analysis strongly suggests that withdrawing VM is likely to cause more human health harm than it prevents, uncertainty about the size of the product $(\Delta F) \times (P^+ - P^-)$ precludes a deterministic conclusion.

6. EXTENSIONS TO CATTLE AND SWINE

To extend the risk assessment in Table 1 to include cattle and swine, it is necessary to compare streptogramin resistance levels in VREF isolates from cattle and pigs to those from chickens. Jensen *et al.* (2002) reported a ratio of streptogramin-resistance in isolates from pigs *vs.* broilers of (0.51/0.67) and a ratio of streptogramin A resistance (which is necessary for QD resistance) of (0.14/0.96). We estimate that the ratio of per-capita consumptions of high-risk (e.g., ground) pork meat to high-risk (e.g., fresh) chicken meat as not more than 0.25. Willems et al (2000) found that only 4 of 87 (4.6%) hospitalized patients had VREFs from the same genogroup as pig VREF isolates, compared to 12% for chicken VREF in Table 1. Thus, even if all QD-resistant VanA VREF from pigs are attributed to VM use, the total QD-resistant VanA VREF risk to humans from pigs might be only about $(0.51/0.67) \times (0.14/0.96) \times (0.25) \times (4.6/12) \approx 0.01$ times as great as from chicken (assuming comparable effects of processing and cooking). Similarly, for cattle, Willems et al (2000) found that, among hospitalized patients, most VREF (84%) belonged to a different genogroup from that in most (70%) veal calf isolates. Assuming that QD-resistant VREF are not more prevalent in beef servings than in chicken or turkey servings (Wegener *et al.*, 1997; Hayes *et al.*, 2003), data on consumption rates of undercooked beef (MRC, 1995) suggest that the human health risk due to beef might be at most about 3% of that from chicken.

In summary, including beef and pork is unlikely to increase estimated human health risk of QD-resistant VanA VREF infections due to VM use in food animals by more than about 4% compared to the risk estimated for chicken alone. In reality, the severely or critically ill patients at risk may be relatively unlikely to be exposed to QD-resistant VREF in undercooked meat from any of these sources.

7. COMPARISON TO RISKS IN AUSTRALIA

An advantage of the RRRT framework is that it can readily be applied to estimate risks in one country by adjusting the parameter estimates for another country. To illustrate, the estimated parameter values in Table 1 would be adjusted as follows to estimate VM-associated risks in Australia:

- VREF cases per year ≈ 16 instead of (37483*0.78 ≈ 29000 in the US), due in part to the smaller population size (data in Chapter 6).
- The fraction of VREF cases that are VanA is only about 0.22 in Australia (Turnridge, 2001) rather than about 0.79 as in the US, shown in Table 1.

- The observed fraction of QD-resistant cases among VREF cases in humans in Australia is 0 (Turnidge and Bell, 2002), although isolates of *E. faecium* from both pork and poultry in Australia have high levels of streptogramin resistance – 81% in retail pig meat and 96.6% in chicken carcasses. Applying a conservative Bayesian approach with a uniform prior (mean = 0.5) to the zero observed QD resistance rates for humans yields a beta posterior distribution with a mean of approximately 0.009 (Chapter 6). [Average streptogramin resistance among food animal sources (chicken, pork, and beef) weighted by consumption exceeds 0.5, suggesting an average foodborne fraction of human VREF cases of less than $0.009/0.5 < 0.02$, as compared to 0.17 in Table 1.]

Holding the other estimated parameters in Table 1 constant, the first two of these adjustments indicate that the current health risks from QD use in food animals are only about $(16/29000) \times (0.22/0.79) = 1.5/10{,}000$ as great as in the US, and hence are vanishingly small – fewer than one excess case expected per millennium.

8. RRRT FOR MACROLIDES AND ENROFLOXACIN

The modularity of the RRRT framework allows much of the work required to estimate and document the factors used for one animal drug to be re-used in assessing the human health impacts of other animal drugs. To illustrate this template-based approach to health impact assessment, this section quickly estimates the human health impacts from continued use of two important classes of antibiotics, macrolides and fluoroquinolones, in chickens. These applications complement the virginiamycin case study, insofar as they address health risks from a foodborne pathogen (*Campylobacter*) instead of a commensal (*E. faecium*); and because fluoroquinolones are used in poultry only for therapeutic purposes, rather than for disease prevention and growth promotion.

To quantify the human health impacts – both positive and negative – of macrolide and fluoroquinolone antibiotic use in animals, we will estimate and compare the following two quantities.

- Preventable RISK to human health from continued use of animal antibiotic = expected additional illness-days *caused* per year by increased antibiotic resistance in foodborne pathogens and *preventable* by ceasing use = (expected preventable resistant cases caused per year) × (expected incremental health consequences per case caused) = $[p(1 - s)(P^{-})MN] \times [f \times r \times (Q_r - Q_s)]$. (The notation in this formula is explained below and summarized in Table 3.)

- BENEFIT (risk reduction) to human health from continued use = expected illness-days *prevented* per year by reduced animal bacterial diseases = (expected cases prevented per year) × (expected health consequences per case) = $[\Delta F(P^+ - P^-)] \times MN \times [Q_r - s'(Q_r - Q_s)]$ where $s' = [1 - (1 - p) \times (1 - s)]$ = post-ban susceptible fraction.

Table 3 summarizes the interpretations and estimated values of the model parameters in these formulas, and they are further discussed and explained below. The first five parameters are the same as in Table 2. An intervention such as withdrawing an animal antibiotic is expected to protect human health if and only if RISK > BENEFIT for continued use.

Table 3: Parameters For Human Health Impact Model

Variable	Meaning	Baseline Value/Source
ΔF	Fractional change in chicken servings from ill or high-risk flocks if current use ceases	0.5% = assumed baseline (Table 2)
P^-	Average probability of illness per serving from animals without disease. Includes indirect effects of cross-contamination.	1.3E-5 (Table 2)
$P^+ - P^- = (1 + R)P^-$	Excess probability of illness per serving from ill flocks. (Includes cross-contamination)	1.2E-4 (Table 2)
M	Average number of servings of food commodity ingested per capita-year	38 (Table 2)
N	Number of people in population	292E6 (Table 2)
$1 - s$	Fraction of the cases caused by bacteria in animal meat that are resistant to human antibiotic. (s = current *susceptible* fraction)	Erythromycin: 0.01 Ciprofloxacin: 0.064
p	Preventable resistance fraction = fraction of currently resistant illnesses caused by eating the food commodity that a ban would remove (i.e., make susceptible)	1 (upper bound) 0.3 may be more realistic for enrofloxacin.
Q_s	Average human health harm (e.g., days of illness or QALYs lost) per susceptible case. Interpreted as "severity" of a case.	6 days (Marano *et al.*, 2000)
$Q_r - Q_s$	Average excess illness-days per resistant case failing to respond normally to antibiotic, for severely ill patients; or per untreated case for non-patients	2 days (Estimated upper bound for current clinical practice Ang and Nacham, 2003; Marano *et al.* 2000)
K	Q_r/Q_s = ratio of average clinical severities (e.g., illness-days) for resistant *vs.* susceptible campylobacteriosis cases	1.002 base case estimate; up to 2 in sensitivity analyses
f	Probability that resistant case fails to respond normally to prescribed antibiotic therapy, due to resistance	< 1 (Upper bound)
r	Probability that a resistant case is assigned resisted antibiotic	0.5

The RISK and BENEFIT formulas have the following interpretations for enrofloxacin, a fluoroquinolone currently used to treat fatal respiratory illness (airsacculitis, AS) in chicken flocks. (MN) = (average servings per capita) × (number of people in US) is the number of chicken servings ingested per year. (P^-)MN is the expected number of resulting campylobacteriosis illness cases per year under the *status quo* (no ban). As in Table 2, P^- denotes the current average risk of illness per serving, i.e., the expected number of illnesses caused per serving. It is estimated from the formula: P^- = (total illnesses per year × fraction caused by eating chicken servings)/(total number of servings). Each illness currently has probability (1 – s) of being resistant to the human antibiotic being considered (e.g., ciprofloxacin). However, if the current animal antibiotic use were to cease, a fraction p of these (1 – s)(P^-)MN currently resistant chicken-caused illnesses, called the *preventable resistance fraction*, would be eliminated – or, more accurately, would be replaced with susceptible rather than resistant pathogens. Thus, a ban would prevent a total of [p(1 – s)(P^-)MN] resistant illnesses per year. This gives the first part of the preventable RISK formula.

For health consequences of a ban, suppose that a fraction f of resistant cases experience reduced treatment effectiveness due to resistance if treated with a resisted antibiotic. Let r denote the probability of being treated with a resisted antibiotic. Thus, r reflects screening and prescription practices, while f reflects the risk that resistance creates clinical harm. If the mean health impact is (Q_r – Q_s) additional illness-days (or quality-adjusted life-years (QALYs) lost, etc.) for each such case, then the average additional health harm per case is f × r × (Q_r – Q_s). Multiplying this average consequence-per-case by the expected number of cases gives the complete formula for estimating the human health risk preventable by a ban on the current animal antibiotic use: RISK = [p(1 – s)(P^-)MN] × [f × r × (Q_r – Q_s)].

The formula for BENEFIT of continued use is interpreted as follows. Suppose that a ban would cause an increase ΔF in the fraction of chicken servings from ill (e.g., airsacculitis-positive, (AS^+)) flocks instead of healthy (e.g., airsacculitis-negative (AS^-)) flocks, and that each such serving has an incremental probability (P^+ – P^-) of causing illness. Then the expected change in number of illnesses per year will be [ΔF(P^+ – P^-)]MN. If a fraction s' = [1 – (1 – p) × (1 – s)] of these illnesses are susceptible after the ban has taken effect [reflecting a pre-ban resistant fraction (1 – s) that is reduced by the preventable fraction p when the effects of the ban are fully realized, leaving (1 – p) × (1 – s) as the new post-ban resistant fraction and hence s' = 1 – (1 – p) × (1 – s) as the new susceptible fraction], then the new average health impact per illness will be [s'Q_s + (1 – s')Q_r], which may be rearranged as [Q_r – s'(Q_r – Q_s)]. Thus, the expected human health impact caused by the

fractional increase ΔF in animal illness prevalence if current animal antibiotic use were to cease is:

$$\text{BENEFIT} = [\Delta F(P^{+} - P^{-})]MN \times [Q_r - s'(Q_r - Q_s)]$$

incremental illness-days per year. This is the human health benefit (= human health harm prevented) from continued use of the animal antibiotic (for which ΔF = 0). Introducing the relative risk ratio $R = (P^{+}/P^{-})$ for the ratio of probability of illness per serving from ill (AS^{+}) *vs.* healthy (AS^{-}) flocks, this formula can be written as:

$$\text{BENEFIT} = [\Delta F(R - 1)] \times [(P^{-})MN)] \times [Q_r - s'(Q_r - Q_s)]$$

The ratio of additional illness-days (or other measures of adverse outcomes) per year that would be *caused* by banning a current animal antibiotic use to illness-days per year *prevented* by the ban is:

$$\begin{aligned}\text{BENEFIT/RISK} &= \\ &[\Delta F(R - 1)] \times [(P^{-})MN)] \times [Q_r - s'(Q_r - Q_s)]/[p(1 - s)(P^{-})MNfr(Q_r - Q_s)] \\ &= [\Delta F(R - 1)][Q_r - s'(Q_r - Q_s)]/[p(1 - s)fr(Q_r - Q_s)].\end{aligned}$$

If a ban would be completely successful in preventing resistance in animals (i.e., the preventable resistance fraction is p = 1, which implies s' = 1), then this BENEFIT:RISK ratio for continued use simplifies to:

$$\text{BENEFIT/RISK} = [\Delta F(R - 1)]/[(1 - s)fr(K - 1)],$$

where $K = (Q_r / Q_s)$ is the ratio of average harm per resistant illness case to average harm per susceptible illness case. "Harm" may be measured in terms of QALYs lost, or the BENEFIT:RISK ratio can be calculated separately for each type of outcome, e.g., mild, moderate, severe, and fatal cases (Buzby *et al.*, 1996.) A ban is health-protective if and only if the BENEFIT:RISK ratio is less than 1.

Henceforth, we will conservatively assume that a ban on current animal antibiotic uses would eliminate *all* resistance in the corresponding food animal-borne resistant campylobacteriosis cases (i.e., p =1 and s' = 1). Thus, we will focus on quantifying the parameters in the following formulas:

$$\begin{aligned}\text{BENEFIT} &= [\Delta F \times (R - 1)] \times [(P^{-}) \times M \times N)] \times Q_s \\ \text{RISK} &= (1 - s) \times [f \times r \times (K - 1)] \times [(P^{-}) \times M \times N] \times Q_s \\ \text{BENEFIT/RISK} &= [\Delta F \times (R - 1)]/[(1 - s) \times f \times r \times (K - 1)]\end{aligned}$$

Recall that P^- denotes the average risk of a campylobacteriosis illness case per chicken serving from a healthy flock (including possible effects of cross-contamination to other foods in the kitchen). Since almost all chicken-borne *C. jejuni* cases currently come from healthy (e.g., AS^-) flocks, P^- can be approximated by dividing the estimated total number of chicken-caused campylobacteriosis cases per year by the total estimated number of chicken servings ingested per year. As discussed for Table 2, the result are:

> P^- = (total chicken-caused cases)/(total servings) = (total cases × fraction from chicken)/(MN) = (1.48E6 × 0.10)/(38 × 292000000) = **1.3E-5** average campylobacteriosis cases caused per chicken serving from a healthy flock.

The corresponding estimate of current cases per year is: $(P^-)MN$ = (1.48E6 × 0.10) = 1.48E5. These estimates are based on population averages, without accounting for interindividual variability in numbers of meals eaten, thoroughness of cooking, differences in immune status and vulnerability, etc. Thus, they should only be used to estimate population risks rather than risks to any specific individual.

As in Table 2, if a linear no-threshold dose-response model is used (i.e., human campylobacteriosis risk is proportional to CFUs per processed carcass) and if the average risk of campylobacteriosis is about **R = 10** times greater for servings from AS^+ flocks compared to those from AS^- flocks (Russell, 2003), then $(P^+ - P^-) = (R - 1) \times (P^-) = 9P^-$ = **1.2E-4** is the excess individual risk of campylobacteriosis per serving from an AS^+ bird. Fitting a log-exponential model to the same data (see Appendix) leads to an estimated R value close to R = 140 instead of R = 10. We use 10 as a conservative baseline estimate, i.e., to reduce the estimated human health benefits of continued animal drug use compared to the estimated benefits of a ban.

For airsacculitis (AS), historical treatment rates with enrofloxacin in the US have been between 0% and 2%. Table 3 assumes that if enrofloxacin and/or macrolides were withdrawn, then AS, NE (Brennan *et al.*, 2001), and perhaps other illnesses that lead to similarly increased microbial loads in processed carcasses (Dawe, 2004) would increase by half a percent, i.e., **ΔF = 0.005**. This is based on assuming that (a) No dramatic increase in flock illness rates (as occurred for NE in Norway) would happen in the United States; but (b) The historical need to treat at least 1% of flocks would continue; and (c) About half of these flocks would be treated successfully by alternatives to enrofloxacin, with the rest being ill at slaughter. (In reality, there may be no fully adequate substitute for enrofloxacin to treat airsacculitis, but perhaps other preventive and therapeutic measures might be developed.) Because this estimate of ΔF is uncertain, the resulting benefits

estimated using Table 3 are benefits *per half-percent increase* in ill flocks following a ban.

The human health consequences of susceptible and resistant campylobacteriosis illnesses are estimated as follows. Although the fractions of chicken-caused severe *C. jejuni* illnesses that are resistant to different antibiotics have not been well studied in the United States, about 1% of all *C. jejuni* illnesses were reported as being erythromycin-resistant in 2000, and this number has been relatively stable or declining for years (CDC, 2000). Thus, Table 3 assumes (1 – s) = 1% among *C. jejuni* for macrolides. The corresponding ciprofloxacin resistance fraction for domestically-acquired *C. jejuni* cases estimated from CDC FoodNet *Campylobacter* Case Control Study data is (1 – s) = 6.4%, (Cox, 2001, pp 110-111). This fraction also appears to be fairly stable over time.

A susceptible case of domestically-acquired chicken-borne campylobacteriosis is assumed to have an average adverse health impact of Q_S = 6 illness-days (Marano *et al.*, 2000). If resistant domestically-acquired chicken-borne campylobacteriosis cases have the same average clinical impacts as susceptible ones, then K = 1 and preventable RISK = 0. This may be the case for adults, as the difference in clinical outcomes between susceptible and resistant domestically-acquired cases is not clear (Ang and Nacham, 2003). However, for purposes of analysis, the base case in Table 3 assumes that resistance leads to an average two-day delay in finding an effective therapy among those cases of campylobacteriosis that are severe enough to warrant antibiotic treatment – about 0.6% of all cases according to Buzby *et al.* (1996) – and that are initially prescribed the resisted antibiotic. Then $K = (Q_r/Q_s) = (0.006 \times 8 \text{ days} + 0.994 \times 6 \text{ days})/(6 \text{ days}) = 1.002$ in the base case (increased to 2 in some of the sensitivity analyses in Table 4), assuming that this delay is the only clinical adverse effect of resistance; that resistance does not otherwise impair recovery (Piddock, 1999); that *all* severe resistant cases are prescribed the resisted antibiotic, and that only those patients with exceptionally severe cases (for which antibiotic treatment might be indicated) are at risk of experiencing a delay in resolution of symptoms from time lost in finding an effective treatment due to resistance.

Even for such patients, excess illness-days occur only of the resisted antibiotic is initially prescribed and then proves ineffective. ["Resistance" of *Campylobacter* to fluoroquinolones refers to an *in vitro* test result that does not necessarily imply clinical resistance (Piddock, 1999).] The probability of being prescribed the resisted antibiotic (e.g., ciprofloxacin for someone with fluoroquinolone-resistant campylobacteriosis, or a macrolide for someone with macrolide-resistant campylobacteriosis) is assumed to be at most r = 0.5 (FDA-CVM, 2001), especially if severe cases are screened for resistance (Ang and Nacham, 2003). The probability that prescription of a resisted

antibiotic leads to compromised treatment or to a treatment failure, thus requiring a switch to a different antibiotic, is uncertain. [One study suggests 1/39 as a possible value (Piddock, 1999)]. It is estimated conservatively in Table 3 as f = 1.

Table 4 calculates the human health RISK and BENEFIT from continued use of enrofloxacin and macrolides in chicken, measured in units of expected illness-days per year caused and prevented, respectively, as well as their RATIO. In addition to the base case values (RATIO = 703 illness-days prevented per illness-day caused for enrofloxacin and 4500 for macrolides), the value of RATIO is calculated for different combinations of the input parameter values, to illustrate the sensitivity of the BENEFIT:RISK ratio to changes in the parameter values.

Table 4: Human Health Impacts of Macrolides and Fluoroquinolones

Input and Meaning	**Base Case**	**Sensitivity Analyses**				
$[(P^{-})MN)]*Q_s$ = current illness-days per year from chicken	8.9E5 = 1.48E5 cases per yr. × 6 days/case					
ΔF = fractional increase in servings from ill flocks if ban	0.005				0.1	0.1
R = Ratio of risk-per-serving from ill *vs.* well flocks	10	2			139	139
(1 – s) = Resistant fraction if no ban	Macrolides: 0.01 Fluoroquinolones: 0.064					
f × r = Adverse clinical outcome probability for resistant cases	0.5 (= prob. given resisted antibiotic)				(0.5)× (1/39)	(0.5)× (1/39)
K = Consequence ratio of illness-days for resistant *vs.* susceptible cases	1.002		1.3	2	2	
Output						
BENEFIT = Illness-days per year *prevented* by continued use	40050 = $[\Delta F(R - 1)] \times [(P^{-})MN)] * Q_s$					
RISK = Illness-days per year *caused* by continued use	57 for enrofloxacin, 9 for macrolides = $(1 - s) \times [fr(K - 1)] \times [(P^{-})MN] \times Q_s$					
RATIO for enrofloxacin = BENEFIT/RISK	703 for enrofloxacin	78	4.7	1.4	1.7E4	8.4E6
RATIO for macrolides	4500 for macrolides	500	30	9	1.1E5	5.4E7

The main conclusion from the baseline calculations is that withdrawing either antibiotic from use in chickens in the US is estimated to cause significantly more illness-days (and more cases of each type of illness, both resistant and susceptible) than it would prevent. The sensitivity analysis columns show how the results change as inputs are varied. (Each column for a sensitivity analysis shows the input values and resulting output values that deviate from the baseline values.) While the health BENEFIT:RISK ratio is only 703 for enrofloxacin and 4500 for macrolides in the base case, which made several conservative assumptions that tend to minimize it, it could be as high as 7 orders of magnitude under pessimistic assumptions (right-most column) in which a ban leads to a 10% increase in chicken flock illness rates; the effects on human illness rates are described by a log-exponential model instead of a linear no-threshold model; and resistance to the prescribed antibiotic has little impact on clinical outcomes (with only 1/39 of patients experiencing excess illness-days, Piddock, 1999.)

Analytic sensitivity analysis is facilitated by the product forms of the RISK and BENEFIT formulas. The fractional change in human health risks caused by a ban is given by $[\Delta F \times (R - 1)]$, and hence will be zero if either ΔF or $(R - 1)$ is zero. As was also the case for virginiamycin, current scientific knowledge does not preclude this possibility, i.e., a ban on macrolides or enrofloxacin in the United States might conceivably turn out to cause no additional animal illnesses and/or no increase in average human illnesses per serving. However, based on the range of values considered in Table 4 and on the European experience and United States data motivating the baseline values, it appears likely that such bans would do far more harm than good to human health. This possibility suggests a high value of information (VoI) for studies directed at clarifying the magnitudes of ΔF and $(R - 1)$ prior to any decision to ban the antibiotics from use in animals.

9. DISCUSSION

9.1 Comparison to Other Risk Assessment Approaches

In contrast to farm-to-fork simulation models, the RRRT approach illustrated in this chapter makes no attempt to identify explicitly all major pathways or mechanisms leading from antibiotic use in food animals to antibiotic resistance in human bacterial infections. Rather, it starts with an observed data point (the number of cases per year in the human population) and works backward to calculate an estimated upper bound on the fraction that might be prevented by removing antibiotic uses in food animals. Table 1 summarizes the fractions used in the calculations for VM. Uncertainties

about the correct values of these fractions (and about the causal pathways and biological phenomena involved) were handled by an upper-bounding approach and through sensitivity analyses. The extent of the biases introduced by repeated use of conservative and/or incorrect assumptions for parameter values in Table 1 is bounded by the fact that none of the fractions can exceed 1.

Uncertainty about future values of the number of cases per year in the human population was examined for VM in Chapter 7, using a population dynamics model of the emergence of resistance that includes the possibilities of colonization, secondary amplification, and person-to-person spread of resistant *E. faecium*. That analysis concluded that the endemic level of resistance in the human population is extremely unlikely to increase as a result of VM use (as the basic reproductive rate R_0 of resistant VREF is much smaller than 1.) In addition, sensitivity analysis of Table 1 indicates that even increasing the number of $VREF_A$ cases per year and/or the QD resistance rate per case 10-fold would not reverse the main finding that the expected human health benefits from continued VM use are much larger than the expected human health risks.

9.2 Comparison of Model Predictions to Experience

Documented Danish experience following the withdrawal of growth promoters provides an opportunity to compare model-predicted human health impacts to observed data. From 1997 to 1998 (when antimicrobial growth promoters were banned) the number of cases of campylobacteriosis in Denmark increased 26.5%, from 2666 to 3372, (Dansk, 1999), while poultry production in Denmark increased by only about 7.4%. The unaccounted-for increase in cases from 1997 to 1998 [$509 = (26.5\% - 7.4\%) \times 2666$] is roughly consistent with the previously estimated 1.35% increase in NE-related contamination in Denmark (detailed under Calculation of ΔF above) and baseline estimate of a 4.5% increase in human campylobacteriosis cases per 0.5% increase in animal illness rates: $(0.045) \times (1.35/0.5) \times 2666 = 324$. In addition, Denmark determined the serotypes of *Campylobacter* infecting humans, broilers, cattle, retail poultry, and healthy dogs (*ibid*). The covariance of serotypes between humans and retail poultry increased from 1997 to 1998, consistent with the hypothesis that increases in *Campylobacter* may have been due to increased contamination from chickens. Finally, the added cases occur in higher age groups, while the campylobacteriosis case rate per 100,000 declined among infants less than 1 year old, consistent with a food source not consumed by infants (e.g. fresh chicken) (Dansk, 1999). In summary, while these sources of evidence are only circumstantial and many other possible historical influences may also have had important effects, the

observed increases in campylobacteriosis illness rates in Denmark are of the same order of magnitude (although somewhat larger than) the impacts predicted by our baseline model.

9.3 Other Considerations and Extensions

In addition to direct effects on microbial loads and resistance fractions of pathogens reaching consumers, a ban on animal antibiotics may have important indirect effects that depend on how decision-makers (e.g., farmers and physicians) adapt to the ban or that are transmitted via causal pathways not addressed in the model. Examples of such additional considerations, with brief comments, are as follows.

- *Antibiotic substitutions and synergies.* After the ban on animal antibiotic growth promoters and prophylactics in Europe, therapeutic use of other animal antibiotics to treat animal diseases increased significantly (Casewell *et al.*, 2003). Similarly, withdrawing virginiamycin use in the United States might cause an increase in ill flocks that could be partly offset by increasing veterinary prescriptions of macrolides and/or other drugs. Similarly, withdrawing enrofloxacin might be compensated for by increasing use of macrolides. Withdrawing all of these antibiotics could increase AS^+ and NE^+ flocks (ΔF in the model) by more than the sum of the increases if each one alone were withdrawn, as compensation with the others would then not be possible.
- *Other animal bacterial diseases.* In addition to airsacculitis and necrotic enteritis, failure to use animal antibiotics might increase the prevalence of other animal illnesses with similar effects on microbial loads in processed chickens (Dawe, 2004). This would increase ΔF.
- *Other foodborne human pathogens.* The preceding assessments have focused on *C. jejuni*. Although *C. coli* cases are only a small percentage of total campylobacteriosis cases, they have much higher resistance rates, e.g., 22.5% against erythromycin, compared to 0.5% for *C. jejuni*, according to Fedorka-Cray et al, 2001. If resistance rates in *C. coli* are about 45 times as great as for *C. jejuni* and *C. coli* constitute a few percent of the total cases, then the human health benefits from withdrawing macrolides could be about double those estimated in Table 4 for *C. jejuni*. Conversely, Russell (2003) reported significant increases (although not on every replicate) of *Salmonella* and *E. coli*, as well as *C. jejuni*, in processed carcasses from AS^+ compared to AS^- flocks. Considering other pathogens might significantly increase the estimated human health benefits from continued use of animal antibiotics.
- *Co-selection and commensals.* Macrolides used in chickens may co-select *E. faecium* that are resistant to streptogramins, although the genes

responsible for resistance to streptogramin A are rarely found in animal isolates. (Since glycopeptides and linezolid are not used in food animals in the US, co-selection risks for these human-use antibiotics are minimal.)

- *Reduced need to treat human patients with antibiotics.* If a ban on VM increases campylobacteriosis cases per year, some of these cases might receive treatment with ciprofloxacin or macrolide antibiotics as empiric treatments. Preventing these cases would remove these human antibiotic prescriptions, potentially reducing selection pressure for resistance in human pathogens and commensals. According to Table 4, a ban or risk management intervention restricting continued use of enrofloxacin could increase total campylobacteriosis by thousands of cases per year. Preventing these cases would remove these human antibiotic prescriptions, potentially reducing selection pressure for resistance in human bacterial pathogens and commensals.
- *Changes in prescription practices.* As physicians and scientists become more concerned about not prescribing antibiotics with doubtful clinical benefits, more rapid diagnostic and resistance-screening tests will be developed and routinely used (e.g., Endtz *et al.*, 2000). These can reduce the prescription rate r for resisted antibiotics in our model, reducing the potential benefits of a ban and increasing RATIO even further.
- *Opportunistic infections and patient practices.* If people treated with antibiotics for other (non-campylobacteriosis) reasons are thereby made significantly more vulnerable to opportunistic infection by antibiotic-resistant *Campylobacter* ingested in chicken or other foods, then the benefits of a ban might be understated in Table 4. Increasing care by patients and at-risk individuals in food preparation, cooking, and handling would tend to attenuate the benefit from this hypothesized source. Empirically, available data (e.g., the raw data of Friedman *et al.*, 2004) do not support the hypothesis that opportunistic resistant infections from chicken-borne *Campylobacter* play a detectable role in human health in the surveyed population.
- *Emergence of resistance.* A common concern is that the resistance fraction (1 – s) may increase over time unless animal antibiotic use is curbed now (APUA, 2002). However, the biomathematical modeling in Chapter 7 suggests that, at least for antibiotics like VM and macrolides that have been used for several decades in food animals without leading to high levels of resistance in people, an outbreak of resistance in the future caused by continued use is very unlikely. Empirically, resistance to macrolides and fluoroquinolones in domestically acquired cases in the United States appears to be fairly stable over time at the levels (about 1% and 6.4%, respectively) in Table 4.

- *Timing:* For simplicity, and to be conservative (i.e., minimizing the BENEFIT:RISK ratio by maximizing the estimated risk of continued use of the animal antibiotic) the *timing* of human health impacts of a ban has so far been ignored: only the new levels that will eventually be reached have been considered. European experience suggests that the hypothesized health benefits to human patients from banning animal antibiotics may take longer than 5 years to materialize (e.g., Heuer *et al.*, 2002; Borgen *et al.*, 2000**,** Iversen *et al.*, 2002), while adverse impacts on increased animal pathogen loads (e.g., Madsen and Pederson, 2000) and possibly on human health (*Eurosurveillance, 2000*) may be much more immediate. If so, then modeling the timing of impacts might further increase the BENEFIT:RISK ratio of continued use of animal antibiotics. Similarly, all of the analyses in Table 4 assume that a ban on animal antibiotic use would be completely effective in eliminating resistance in servings from food animals (i.e., p = 1). In reality, experience in Europe, as well as in Canada after reduced fluoroquinolone use (comparing resistance data in Gaudreau and Gilbert 2003 *vs.* 1998), suggests that the preventable resistance fraction may be much less than 1, tending to increase the BENEFIT:RISK ratio for continued use.

In summary, while the analyses in Tables 1-4 have focused on QD-resistant VanA VREF and on *Campylobacter* illnesses transmitted via chicken servings, other important considerations and extensions to consider other pathogens may tend to strengthen the conclusion that human health risks from withdrawing or restricting animal antibiotics could significantly outweigh potential human health benefits.

Such additional comparisons and information can be included in expanded quantitative human health risk assessments. However, doing so may have limited value from a decision analysis point of view if the main effect is to further strengthen the already strong conclusion from the analyses in Tables 1-4. A key prescriptive principle of value-of-information (VoI) analysis is not to pay for information that does not have the potential to change risk management decision. By contrast, better information on the extent to which withdrawing animal antibiotics increases animal disease rates and microbial loads in animal carcasses and resulting illness risks to consumers (the size of $(\Delta F) \times (P^{+} - P^{-})$ in the model) could be very valuable in reducing uncertainty about the baseline conclusion that continued animal antibiotic use has human health benefits that are likely to be far larger than its human health risks.

10. SUMMARY AND CONCLUSIONS

This chapter has illustrated applications of the Rapid Risk Rating Technique (RRRT) approach for estimating the human health impacts of animal antibiotic uses. The approach appears to be practical to implement with available data for antibiotics and pathogens of practical interest. Potential human health benefits from discontinuing animal antibiotic uses are estimated by multiplying total clinical case rates by a sequence of fractions estimated from data (or conservatively bounded, e.g., by setting highly uncertain fractions equal to their maximum possible value of 1) to estimate the number of potentially preventable cases and adverse consequences per year. Potential human health risks from discontinuing use are estimated by multiplying the expected increase in food animal illness rates by the estimated increase in human illnesses (and resulting adverse consequences) per year per unit increase in animal illness rates.

This approach is designed for use in situations in which qualitative considerations are insufficient to support clear, effective risk management decision-making. This may occur when the quantitative sizes of health risks matter, or if qualitative assessment indicates that a proposed intervention may cause both human health benefits and human health risks, and it is important to determine which is likely to be larger. The approach is also appropriate for situations in which available data and resources are not sufficient to build and validate a more detailed quantitative model, such as the one in Chapter 7, or when rapid approximate risk assessment using the RRRT formulas provides a sufficiently clear answer so that more expensive and detailed quantitative estimates are unnecessary. The risk and benefit factors estimated in Tables 1-4, are intended to provide the least amount of information that is both necessary and sufficient to estimate and compare quantitative human health risks and benefits from alternative risk management interventions.

In summary, the approach to human health impact assessment developed and illustrated in this chapter is intended to deliver the major benefits that qualitative risk assessment approaches have sought to provide, outlined in Chapter 1 – such as practicality of input data requirements, clear logic and calculations (using basic arithmetic operations to calculate human health risks and benefits from parameters estimated from available data), and easily interpretable outputs (does RISK exceed BENEFIT?) – while preserving the advantages of soundness (assigning larger risk estimates to larger risks) and ability to deal with uncertainties (through bounds and intervals) made possible by quantitative assessment methods. The approach was illustrated using virginiamycin in chickens as a case study, with fluoroquinolones and macrolides as additional examples.

The baseline estimates suggest that the human health harm from discontinuing use of these antibiotics in food animals may substantially exceed the potential human health benefits. Scientific uncertainties preclude available data from proving with certainty which option (i.e., continuing or discontinuing antibiotic use) will create fewer adverse human health impacts. However, the baseline calculations and sensitivity analyses strongly suggest that continued use of VM and other antibiotics would prevent at least thousands of times more illness-days (as well as more fatalities and cases) per year than it would cause. Thus, any risk management policies that advocate withdrawing these animal antibiotic uses for precautionary or other motives must be based on principles, assumptions, or data very different from those considered in this chapter.

The analyses undertaken in this chapter strongly suggest that the most prudent present risk management policy for several classes of animal antibiotics – including streptogramins, macrolides, fluoroquinolones, and other antibiotics with similar effects in animal and human medicine – is to *continue to use them to help prevent animal and human illnesses,* rather than using them exclusively in treating human illnesses. To reverse this conclusion and make a ban on animal antibiotics a rationally recommended strategy, it would be necessary to show that animal antibiotic uses are much less effective in preventing animal illnesses (or other "high risk" conditions) than was assumed in Tables 1-4, or that servings from ill animals have much less impact on human health than was estimated in those tables.

If the analysis presented here is indeed conservative, as intended, then collecting additional data to reduce current scientific uncertainties will be more likely to strengthen than to undermine or refute the conclusion that using antibiotics to help *prevent* animal (and hence human) illnesses protects human health much more effectively than reserving antibiotics for use in *treating* human illnesses. But this is only a best bet, given what is known now. The techniques illustrated in this chapter for practical quantitative risk assessment and comparison of likely human health consequences of alternative risk management interventions can be applied in future, as additional scientific data are collected, to narrow present uncertainty bounds and to reach more definitive conclusions that reflect growing understanding of the causal links between animal and human health.

APPENDIX: THE LOG-EXPONENTIAL RISK MODEL

FDA's Center for Veterinary Medicine (FDA-CVM, 2001, www.fda.gov/cvm/antimicrobial/RRAIntro.pdf) suggested a log-exponential model for risk assessment of *Campylobacter*. In this model, the log of the microbial load of *Campylobacter* reaching consumers via chicken servings has an exponential distribution, placing greater probability densities on smaller microbial loads. Microbial loads at ingestion are assumed to be proportional to (but smaller than) microbial loads on carcasses following processing. An uncertain reduction factor α expresses the proportionality between CFU/ml measured in processing rinse fluids and CFU per chicken serving ingested. These assumptions imply the mathematical model:
$\Pr(\log \text{ingested dose} > x) = e^{-\alpha\lambda x}$, where $1/\lambda$ = mean number of log-CFU/ml of chicken rinse fluid at post-processing ($= \ln(10^{1.09}) = 2.51$ based on data of Russell, 2003, implying an estimated value for λ of: $\lambda = 1/2.51 = 0.3984$.) Let MID denote the minimum infectious dose (possibly 1 CFU). Substituting log(MID) for x then gives the following formula:

$$\Pr(\text{ingested dose} > \text{MID}) = \Pr[\log(\text{ingested dose}) > \log(\text{MID})] = e^{-\alpha\lambda\log(\text{MID})}.$$

Independently, the probability of an infectious dose in a chicken serving under current conditions (i.e., Pr(ingested dose > MID)) can also be estimated from data, as follows:

- *Main formula:* Pr(ingested dose in a chicken serving > MID) = (total cases per year from chicken servings)/(total chicken servings per year * fraction of infectious servings that cause illness)
- *Total cases per year from chicken consumption* ≈ 1.48E5 cases per year. (This is $(P^{-})MN$ in Table 2.)
- *Total chicken servings per year* = MN = (38 fresh chicken servings/capita-year)*(292E6), based on the data in Table 2.
- *Fraction of infectious servings that cause illnesses* was estimated by Rosenquist *et al.*, 2003 as 0.22 (11/50) with a beta uncertainty distribution. WHO, 2002 states: "Overall, pooling all the data, a total of 29 people got sick out of 89 individuals that were infected (33%)." Finch and Blake (1985) report a median attack rate of **0.41** in outbreaks following high exposures in various food vehicles. We will use **0.41** as it is based on a mix of real populations, strains, and food vehicles.

Substituting the above values into the main formula gives:

$\Pr(\text{ingested dose} > \text{MID} \mid AS^{-}) = (1.48E5 \text{ cases/year})/(38*292E6*0.41) =$ **3.23E-5.**

Equating this empirical estimate to the above model-based formula then gives: Pr(ingested dose > MID | AS^-) = $e^{-\alpha\lambda\log(MID)}$ = 3.23E-5, which rearranged, becomes: $\alpha\lambda$*log(MID) = -ln(3.23E-5) = 10.34, or, on substituting the estimate λ = 0.3984: α*log(MID) = 10.34/λ = 10.34/0.3984 = 25.954.

Now suppose that λ changes to reflect an average tenfold increase in microbial load per processed carcass, from $\lambda = 1/\ln(10^{1.09}) = 0.3984$ for AS^- flocks to $\lambda^+ = 1/\ln(10^{2.09}) = 0.2078$ (based on the data of Russell, 2003). Then risk per serving will become:

Pr(ingested dose > MID | AS+) = exp[-α*log(MID)*λ^+]
= exp(-25.954*0.2078) = 0.0045.

The relative risk of human illness due to untreated airsacculitis in chickens is therefore:

R = Pr(exposure > MID | AS^+)/Pr(exposure > MID | AS^-) = 0.0045/3.23E-5 = **139.3**.

R does not depend on the assumed dose-response threshold MID (the effective minimum infective dose, if any), due to a well-known property of exponential distributions. Thus, the increase in human health risk if a fraction ΔF of currently AS^- flocks become AS^+ is:

Risk from ΔF = ΔF*(R – 1) = 138.3*ΔF.

This page intentionally blank

References

Aarestrup FM, Seyfarth AM, Emborg HD, Pedersen K, Hendriksen RS, Bager F. Effect of abolishment of the use of antimicrobial agents for growth promotion on occurrence of antimicrobial resistance in fecal enterococci from food animals in Denmark. Antimicrob Agents Chemother. 2001 Jul;45(7):2054-9.

Acar JF, Goldstein FW. Trends in bacterial resistance to fluoroquinolones. Clin Infect Dis. 1997 Jan;24 Suppl 1:S67-73.

Adak GK, Cowden JM, Nicholas S, Evans HS. The Public Health Laboratory Service national case-control study of primary indigenous sporadic cases of *Campylobacter* infection. Epidemiol Infect. 1995 Aug;115(1):15-22.

Afset JE, Maeland JA. [Erythromycin and ciprofloxacin resistant *Campylobacter* jejuni] Tidsskr Nor Laegeforen. 2001 Aug 10;121(18):2152-4. Norwegian.

AHA. 2001. *Hospital Statistics*. Health Forum – American Hospital Association. (http://www.aha.org/resource).

Ahn WK, Bailenson J. Causal attribution as a search for underlying mechanisms: an explanation of the conjunction fallacy and the discounting principle. Cognit Psychol. 1996 Aug;31(1):82-123.

Aliferis C, Tsamardinos I, Statnikov A. 2003. HITON: a novel Markov Blanket algorithm for optimal variable selection.AMIA Annu Symp Proc. 2003;:21-5.

Alliance for Prudent Use of Antibiotics (APUA): FAAIR Report, The Need to Improve Antimicrobial Use in Agriculture: Ecological and Human Health Effects. *Clinical Infectious Diseases*, Vol 34, Supplement 3, June 1, 2002. http://www.healthsci.tufts.edu/apua/Ecology/faair.html

Allos BM. *Campylobacter* jejuni Infections: update on emerging issues and trends. Clin Infect Dis. 2001 Apr 15;32(8):1201-6. Epub 2001 Mar 28.

Altekruse SF, Stern NJ, Fields PI, Swerdlow DL. Campylobacter jejuni – an emerging foodborne pathogen.Emerg Infect Dis. 1999 Jan-Feb;5(1):28-35.

AMR. 2001. U.S. Hospital Anti-Infective Market Guide. AMR, Inc., Malvern, PA. (http://www.amr-data.com).

Anders BJ, Lauer BA, Paisley JW, Reller LB. Double-blind placebo controlled trial of erythromycin for treatment of *Campylobacter* enteritis. Lancet. 1982 Jan 16;1(8264):131-2.

Andrews GP. The enteric Campylobacter: they are everywhere. Clin Lab Sci. 1998 Sep-Oct;11(5):305-8.

Andrieu C, de Freitas N, Doucet A, and Jordan MI. An introduction to MCMC for machine learning, Machine Learning, vol. 50, pp. 5--43, Jan. - Feb. 2003. http://citeseer.csail.mit.edu/andrieu03introduction.html

Ang JY, Nachman S. *Campylobacter* Infections. http://www.emedicine.com/ped/topic2697.htm

Angulo FJ, Nargund VN, Chiller TC. Evidence of an association between use of anti-microbial agents in food animals and anti-microbial resistance among bacteria isolated from humans and the human health consequences of such resistance. J Vet Med B Infect Dis Vet Public Health. 2004 Oct-Nov;51(8-9):374-9.

APUA (Alliance for Prudent Use of Antibiotics): FAAIR Report, The Need to Improve Antimicrobial Use in Agriculture: Ecological and Human Health Effects. Clinical Infectious Diseases, Vol 34, Supplement 3, June 1, 2002. http://www.healthsci.tufts.edu/apua/Ecology/faair.html.

Austin, DJ ,MJM Bonten, RA Weinstein, S Slaughter, and RM Anderson. 1999. Vancomycin-resistant enterococci in intensive-care hospital settings: Transmission dynamics, persistence, and the impact of infection control programs. *Proc. Natl. Acad. Sci. USA* 96, 6908–6913. (http://www.pnas.org/).

Austin PC, Tu JV. Automated variable selection methods for logistic regression produced unstable models for predicting acute myocardial infarction mortality. J Clin Epidemiol. 2004 Nov;57(11):1138-46.

Avcare (2003). The Role of Enteric Antibiotics in Livestock Production. http://www.avcare.org.au/files/animalhealth/information/The%20Role%20of%20enteric%20antibiotics%20in%20livestock%20production.pdf.

Aventis, 2002. Aventis Annual Report, 2001.

http://www.aventis.com/main/order_center/download/ave_annualreport_2001_short_en.pdf

Bafundo KW, Cox LA, Jr., Bywater R, 2003. The use of virginiamycin in food animal production. *Feedstuffs.* Jan. 20, 2003. 26-27.

Bager F, Aarestrup FM, Madsen M, Wegener HC. Glycopeptide resistance in Enterococcus faecium from broilers and pigs following discontinued use of avoparcin. Microb Drug Resist. 1999 Spring;5(1):53-6.

Bagley SC, White H, Golomb BA. Logistic regression in the medical literature: standards for use and reporting, with particular attention to one medical domain. J Clin Epidemiol. 2001 Oct;54(10):979-85.

Bailar JC 3rd, Travers K. Review of assessments of the human health risk associated with the use of antimicrobial agents in agriculture. Clin Infect Dis. 2002 Jun 1;34 Suppl 3:S135-43.

Banffer JR. Biotypes and serotypes of *Campylobacter* jejuni and *Campylobacter* coli strains isolated from patients, pigs and chickens in the region of Rotterdam. J Infect. 1985 May;10(3):277-81.

Barber DA, Miller GY, McNamara PE.. Models of antimicrobial resistance and foodborne illness: examining assumptions and practical applications. J Food Prot. 2003 Apr;66(4):700-9.

Barbour EK, Hamadeh S, Talhouk R, Sakr W, Darwish R. Evaluation of an enrofloxacin-treatment program against Mycoplasma gallisepticum infection in broilers. Prev Vet Med. 1998 May 1;35(2):91-9.

Barbour AD, Mansson M. Compound Poisson approximation. Annals of Probability 2000;30(3):1492-1537.

Bartholomew MJ, Vose DJ, Tollefson LR, Travis CC. A linear model for managing the risk of antimicrobial resistance originating in food animals. Risk Anal. 2005 Feb;25(1):99-108.

Barza M, Travers K. Excess infections due to antimicrobial resistance: the "Attributable Fraction". Clin Infect Dis. 2002 Jun 1;34 Suppl 3:S126-30.
http://www.journals.uchicago.edu/CID/journal/contents/v34nS3.html

Bastardi A, Shafir E.On the pursuit and misuse of useless information. J Pers Soc Psychol. 1998 Jul;75(1):19-32.

Bazan S. Enhancing decision-making effectiveness in problem-solving teams. Clin Lab Manage Rev. 1998 Jul-Aug;12(4):272-6.

Bell J, Turnidge J, Coombs G, O'Brien F. Emergence and epidemiology of vancomycin-resistant enterococci in Australia. Commun Dis Intell. 1998 Oct 29;22(11):249-52.

Benson A, Kaplan JE, Masur H, Pau A, Holmes KK. Treating opportunistic infections among HIV-Infected adults and adolescents: Recommendations from CDC, the National Institutes of Health, and the HIV Medicine Association / Infectious Diseases Society of America. Clinical Infectious Diseases. 2005 40:S131–235. http://www.cdc.gov/mmwr/preview/mmwrhtml/rr5315a1.htm

Bilgic, T. Interval-valued preference structures. 1997. http://www.ie.boun.edu.tr/~taner/publications/papers/ejor.pdf

Bischoff WE, Reynolds TM, Hall GO, Wenzel RP, Edmond MB. Molecular epidemiology of vancomycin-resistant Enterococcus faecium in a large urban hospital over a 5-year period. J Clin Microbiol. 1999 Dec;37(12):3912-6.

Blaine K, Powell D. Communication of food-related risks. AgBioForum . Volume 4, Number 3&4 . 2001 . Pages 179-185
http://www.foodsafetynetwork.ca/food/blainepowell.pdf

Blaser MJ, Sazie E, Williams LP Jr. The influence of immunity on raw milk--associated *Campylobacter* infection. JAMA. 1987 Jan 2;257(1):43-6.

Blaser MJ, LaForce FM, Wilson NA, Wang WL. Reservoirs for human campylobacteriosis. J Infect Dis. 1980 May;141(5):665-9.

Blaser MJ, Checko P, Bopp C, Bruce A, Hughes JM. *Campylobacter* enteritis associated with foodborne transmission. Am J Epidemiol. 1982 Dec;116(6):886-94.

Boerlin P, Wissing A, Aarestrup FM, Frey J, Nicolet J. Antimicrobial growth promoter ban and resistance to macrolides and vancomycin in enterococci from pigs. J Clin Microbiol. 2001 Nov;39(11):4193-5.

Borgen K, Simonsen GS, Sundsfjord A, Wasteson Y, Olsvik O, Kruse H. Continuing high prevalence of VanA-type vancomycin-resistant enterococci on Norwegian poultry farms three years after avoparcin was banned. J Appl Microbiol. 2000 Sep;89(3):478-85.

Bornstein BH, Emler AC. Rationality in medical decision making: a review of the literature on doctors' decision-making biases. J Eval Clin Pract. 2001 May;7(2):97-107.

Brennan J, Moore G, Poe SE, Zimmermann A, Vessie G, Barnum DA, Wilson J. Efficacy of in-feed tylosin phosphate for the treatment of necrotic enteritis in broiler chickens. Poult Sci. 2001 Oct;80(10):1451-4.

Brieseman MA. A further study of the epidemiology of *Campylobacter* jejuni infections.N Z Med J. 1990 May 9;103(889):207-9.

Brown P, Kidd D, Riordan T, Barrell RA. An outbreak of food-borne *Campylobacter* jejuni infection and the possible role of cross-contamination. J Infect. 1988 Sep;17(2):171-6.

Burges CJC. A Tutorial on Support Vector Machines for Pattern Recognition. *Knowledge Discovery and Data Mining*, 2(2), 1998. http://citeseer.ist.psu.edu/42037.html

Burmaster DE, Anderson PD. Principles of good practice for the use of Monte Carlo techniques in human health and ecological risk assessments. Risk Anal. 1994 Aug;14(4):477-81.

Butzler JP. Campylobacter, from obscurity to celebrity. Clin Microbiol Infect. 2004 Oct;10(10):868-76.

Buzby, JC, T Roberts, CT Jordan Lin, and JM MacDonald. 1996. Bacterial Foodborne Disease: Medical Costs and Productivity Losses. USDA, Economic Research Service, Agricultural Economics Report # 741. (http://www.ers.usda.gov/publications/aer741/)

Byrd, DM III, Cothern CR. Introduction to Risk Analysis: A Systematic Approach To Science-Based Decision Making. ABS Consulting. 2000.

Campylobacter sentinel surveillance scheme collaborators. Ethnicity and *Campylobacter* infection: a population-based questionnaire survey. J Infect. 2003 Oct;47(3):210-6.

Cascante M, Franco R, Canela EI. Use of implicit methods from general sensitivity theory to develop a systematic approach to metabolic control. II. Complex systems.Math Biosci. 1989 Jun;94(2):289-309.

Casewell M, Friis C, Marco E, McMullin P, Phillips I. The European ban on growth-promoting antibiotics and emerging consequences for human and animal health. J Antimicrob Chemother. 2003 Aug;52(2):159-61.

Cassin MH, Paoli GM, Lammerding AM. Simulation modeling for microbial risk assessment. J Food Prot. 1998 Nov;61(11):1560-6.

CDC, 2000. www.cdc.gov/narms/annual/2000/annual_pdf.htm

CDC. 2002. FoodNet Surveillance Report for 2000 (Final Report). Centers for Disease Control and Prevention, FoodNet Surveillance Program (http://www.cdc.gov/foodnet).

CDC, 2003. Preliminary FoodNet Data on the Incidence of Foodborne Illnesses --- Selected Sites, United States, 2002. *MMWR*, 52(15), 340-343. http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5215a4.htm)

CDC DBMD Disease Information, http://www.cdc.gov/ncidod/dbmd/diseaseinfo/campylobacter_g.htm

Chaudhuri P, Lo W-D, Loh W-Y, Yang C-C. Generalized Regression Trees. *Statistica Sinica* 1995, v. 5, pp. 641-666. http://www.stat.wisc.edu/~loh/treeprogs/guide/grapes.pdf

Chang KC, Tian Z, 2002. Efficient inference for mixed Bayesian networks. http://citeseer.nj.nec.com/cachedpage/570208/1

Chartier J., Gabler S. Risk Communication and Government: Theory and Application for the Canadian Food Inspection Agency. 2001. http://www.inspection.gc.ca/english/corpaffr/publications/riscomm/riscomme.shtml

Cheng J, Druzdzel. 2000. AIS-BN: An Adaptive Importance Sampling Algorithm for Evidential Reasoning in Large Bayesian Networks. http://citeseer.nj.nec.com/cheng00aisbn.html

Choi BC, Noseworthy AL. Classification, direction, and prevention of bias in epidemiologic research. J Occup Med. 1992 Mar;34(3):265-71.

Christensen B, Sommer H, Rosenquist H, Nielsen N. Risk Assessment on *Campylobacter jejuni* in Chicken Products: First Edition http://www.foedevaredirektoratet.dk/NR/rdonlyres/emmun3ukyfeweov2kt5f72bngkszsy3g4tfwzd2do4qlyn7my2bseotxcrdyknz7bndtoe2p6pkzmszbqabatvo4cje/rapport_jan2001.pdf

Christiansen KJ, Tibbett PA, Beresford W, Pearman JW, Lee RC, Coombs GW, Kay ID, O'Brien FG, Palladino S, Douglas CR, Montgomery PD, Orrell T, Peterson AM, Kosaras FP, Flexman JP, Heath CH, McCullough CA. Eradication of a large outbreak of a single strain of vanB vancomycin-resistant Enterococcus faecium at a major Australian teaching hospital. Infect Control Hosp Epidemiol. 2004 May;25(5):384-90.

Clark, NC, RC Cooksey, BC Hill, JM Swenson, and FC. Tenover. 1993. Characterization of glycopeptide-resistant enterococci from U.S. hospitals. *Antimicrob. Agents Chemother*, 37,2311–2317.

Clemen, RT and Reilly, T. *Making Hard Decisions with Decision Tools Suite.* Duxbury Press. 2000.

Cliff, AD and JK Ord. *Spatial Processes: Models and Applications*. Pion Limited, London, 1981.

Coleman ME, Sandberg S, Anderson SA. Impact of microbial ecology of meat and poultry products on predictions from exposure assessment scenarios for refrigerated storage. Risk Anal. 2003 Feb;23(1):215-28.

Cox LA Jr., *Risk Analysis: Foundations, Models, and Methods.* Kluwer Academic Press, Boston, MA. 2001.

Cox LA Jr. Re-examining the causes of Campylobacteriosis. *International Journal of Infectious Diseases,* December, 2002.

Cox LA Jr. "Mortality associated with foodborne bacterial gastrointestinal infections reexamined**".** *British Medical Journal.* Rapid Response. February 19th, 2003. http://bmj.com/cgi/eletters/326/7385/357#29767

Cox LA Jr. Domestically acquired fluoroquinolone-resistant Campylobacter infection. Clin Infect Dis. 2004 Nov 1;39(9):1399-400; author reply 1400-1.

Cox LA Jr. Potential human health benefits of antibiotics used in food animals: A case study of virginiamycin. Environment International, 2005 May;31(4): 549-563.

Cox LA Jr., Babayev D, Huber W. Some limitations of qualitative risk rating systems. *Risk Analysis,* 2005 Jun;25(3):651-62.

Cox LA Jr., Popken, DA. A simulation model of human health risks from chicken-borne Campylobacter jejuni. Technology 9;55-84, 2002

Cox LA Jr., Popken DA. Quantifying human health risks from virginiamycin used in chickens. Risk Anal. 2004a Feb;24(1):271-88.

Cox LA Jr., Popken DA. Bayesian Monte Carlo uncertainty analysis of human health risks from animal antimicrobial use in a dynamic model of emerging resistance. Risk Anal. 2004b Oct;24(5):1153-64.

Cox LA Jr., Popken DA 2005. Quantifying potential human health impacts of animal antibiotics: Enrofloxacin and macrolides in chickens. Forthcoming in *Risk Analysis.* www.sra.org/news0203.pdf

Cox LA Jr., Ricci PF.. Causation in risk assessment and management: models, inference, biases, and a microbial risk-benefit case study. Environ Int. 2005 Apr;31(3):377-97.

Critchley IA, Blosser-Middleton RS, Jones ME, Thornsberry C, Sahm DF, Karlowsky JA. Baseline study to determine in vitro activities of daptomycin against gram-positive pathogens isolated in the United States in 2000-2001. Antimicrob Agents Chemother. 2003 May;47(5):1689-93.

Cromwell GL. Why and how antibiotics are used in swine production. Anim Biotechnol. 2002 May;13(1):7-27.

DANMAP. 1997-2000. Consumption of antimicrobial agents and occurrence of antimicrobial resistance in bacteria from food animals foods and humans in Denmark http://www.svs.dk/dk/Publikationer/Frm_pub.htm

Dansk. 1999. Annual Report on Zoonoses in Denmark 1998. Dansk Zoonosecenter. http://zoonyt.dzc.dk/annualreport1998/index.html

Dansk. 1998. Annual Report on Zoonoses in Denmark 1997. Dansk Zoonosecenter. http://zoonyt.dzc.dk/annualreport1997/index.html

Darwiche A, Goldszmidt M. On the relation between kappa calculus and probabilistic reasoning. In R. Lopez de Mantaras and D. Poole, editors, Uncertainty in Artificial Intelligence, volume 10, pages 145-153. Morgan Kaufmann, San Francisco, CA, 1994.

Davidson VJ, Ryks J. Comparison of Monte Carlo and fuzzy math simulation methods for quantitative microbial risk assessment. J Food Prot. 2003 Oct;66(10):1900-10.

Dawe, J., 2004. The relationship between poultry health and food safety. The Poultry Informed Professional. 2004 April (77):1-6. http://www.avian.uga.edu/documents/pip/

Dechter, R. 1999. Bucket Elimination: A Unifying Framework for Reasoning. http://citeseer.nj.nec.com/520583.html

Deeks JJ, Dinnes J, D'Amico R, Sowden AJ, Sakarovitch C, Song F, Petticrew M, Altman DG; International Stroke Trial Collaborative Group; European Carotid Surgery Trial Collaborative Group. Evaluating non-randomised intervention studies. Health Technol Assess. 2003;7(27):iii-x, 1-173.

Del Grosso MD, Caprioli A, Chinzari P, Fontana MC, Pezzotti G, Manfrin A, Giannatale ED, Goffredo E, Pantosti A. Detection and characterization of vancomycin-resistant enterococci in farm animals and raw meat products in Italy. Microb Drug Resist. 2000 Winter;6(4):313-8.

Deming MS, Tauxe RV, Blake PA, Dixon SE, Fowler BS, Jones TS, Lockamy EA, Patton CM, Sikes RO. *Campylobacter* enteritis at a university: transmission from eating chicken and from cats. Am J Epidemiol. 1987 Sep;126(3):526-34. Erratum in: Am J Epidemiol 1987 Dec;126(6):1220.

Desmonts MH, Dufour-Gesbert F, Avrain L, Kempf I. Antimicrobial resistance in *Campylobacter* strains isolated from French broilers before and after antimicrobial growth promoter bans. J Antimicrob Chemother. 2004 Dec;54(6):1025-1030.

Donskey CJ, Salata RA, 2003. Enterococcal Infections. http://www.emedicine.com/med/topic680.htm#

DPI. 2000. How Broilers are Marketed, Delmarva Poultry Industry, Inc., http://www.dpichicken.org

Drake, A, J McClellan, K Joyce, T Barrett, F Angulo and the NARMS Enterococci Working Group. 2002. High-level gentamicin-resistant enterococci and quinipristin/dalfopristin-resistant *E. faecium* from ground pork purchased from grocery stores. National Antimicrobial Resistance Monitoring Systems, Annual Scientific Meeting. November 19-22. Hilton Head, SC. (http://www.cdc.gov/narms/pub/presentations/narms/2002/draje_a.htm.)

Druzdzel , MJ, 1994. Some Useful Properties of Probabilistic Knowledge Representations From the Point of View of Intelligent Systems. http://citeseer.nj.nec.com/druzdzel94some.html

Dryden MS, Gabb RJ, Wright SK. Empirical treatment of severe acute community-acquired gastroenteritis with ciprofloxacin. Clin Infect Dis. 1996 Jun;22(6):1019-25.

Eberhart-Phillips J, Walker N, Garrett N, Bell D, Sinclair D, Rainger W, Bates M. Campylobacteriosis in New Zealand: results of a case-control study. J Epidemiol Community Health. 1997 Dec;51(6):686-91.

Edmond MB, Ober JF, Weinbaum DL, Pfaller MA, Hwang T, Sanford MD, Wenzel RP. Vancomycin-resistant Enterococcus faecium bacteremia: risk factors for infection. Clin Infect Dis. 1995 May;20(5):1126-33.

Effler P, Leong MC, Kimura A, *et al.*. Sporadic *Campylobacter* jejuni infections in Hawaii: associations with prior antibiotic use and commercially prepared chicken. J Infect Dis 2001; 183(7):1152–1155.

Eliopoulos, GM, CB Wennersten, HS Gold, T Schulin, M Soull, MG Farris, S Cerwinka, HL Nadler, M Dowzicky, GH Talbot, and RC Moellering, Jr. 1998. Characterization of Vancomycin resistant Enteroccocus faecium isolates from the United States and their susceptibility in-vitro to Dalfopristin-Quinupristin. *Antimicrobial Agents and Chemotherapy*, 42(5), 1088-1092. (http://aac.asm.org)

Elstein AS. Heuristics and biases: selected errors in clinical reasoning. Acad Med. 1999 Jul;74(7):791-4.

Endtz HP, Ang CW, van den Braak N, Luijendijk A, Jacobs BC, de Man P, van Duin JM, van Belkum A, Verbrugh HA. Evaluation of a new commercial immunoassay for rapid detection of *Campylobacter* jejuni in stool samples. Eur J Clin Microbiol Infect Dis. 2000 Oct; 19(10): 794-7.

Endtz HP, Ruijs GJ, van Klingeren B, Jansen WH, van der Reyden T, Mouton RP. Quinolone resistance in *Campylobacter* isolated from man and poultry following the introduction of fluoroquinolones in veterinary medicine. J Antimicrob Chemother. 1991 Feb;27(2):199-208.

Engberg J, Neimann J, Nielsen EM, Aerestrup FM, Fussing V. Quinolone-resistant *Campylobacter* infections: risk factors and clinical consequences. Emerg Infect Dis. 2004 Jun;10(6):1056-63.

Engster, HM, D Marvil, and B Stewart-Brown. 2002. The Effect of Withdrawing Growth Promoting Antibiotics from Broiler Chickens, a Long-Term Commercial Industry Study. *Journal of Applied Poultry Research*, 11, 431-6.

Ericsson CD, Johnson PC, Dupont HL, Morgan DR, Bitsura JA, de la Cabada FJ. Ciprofloxacin or trimethoprim-sulfamethoxazole as initial therapy for travelers' diarrhea. A placebo-controlled, randomized trial. Ann Intern Med. 1987 Feb;106(2):216-20.

Eurosurveillance Weekly. Campylobacteriosis in Norway 2001: incidence still rising. *Eurosurveillance Weekly*, 6(24), 13 June 2002. http://www.eurosurveillance.org/ew/2002/020613.asp http://www.eurosurv.org/2002/rtf/020613.rtf;

FAO/WHO, 2001. Joint FAO/WHO Expert Consultation on Risk Assessment of Microbiological Hazards in Foods: Risk characterization of Salmonella spp. in eggs and broiler chickens and Listeria monocytogenes in ready-to-eat foods. FAO Headquarters, Rome, Italy 30 April - 4 May 2001
http://www.who.int/foodsafety/publications/micro/en/may2001.pdf

FAO/WHO, 2002. Hartnett E, Paoli G, Fazil A, Lammerding A, Anderson S, Rosenquist H, Christensen B, Nuata M. Risk assessment of *Campylobacter* spp. in broiler chickens and Vibrio spp. in seafood. Report of a Joint FAO/WHO Expert Consultation Bangkok, Thailand. 5-9 August 2002. ftp://ftp.fao.org/es/esn/food/cv_02e.pdf

FDA-CFSAN, 2001. U. S. Food and Drug Administration Center for Food Safety and Applied Nutrition (CFSAN). Draft Risk Assessment on the Public Health Impact of *Vibrio parahaemolyticus* in Raw Molluscan Shellfish. January 2001
http://vm.cfsan.fda.gov/~dms/vprisk6.html

FDA-CFSAN, 2003. FDA/Center for Food Safety and Applied Nutrition USDA/Food Safety and Inspection Service Centers for Disease Control and Prevention. Quantitative Assessment of Relative Risk to Public Health from Foodborne *Listeria monocytogenes* Among Selected Categories of Ready-to-Eat Foods. Appendix 3. September 2003. http://vm.cfsan.fda.gov/~dms/lmr2-a3.html

FDA-CVM. 2001. "Human Health Impact of Fluoroquinolone Resistant *Campylobacter* Attributed to the Consumption of Chicken", US Food and Drug Administration, Center for Veterinary Medicine (Revised Jan, 2001). http://www.fda.gov/cvm/antimicrobial/RRAsec5.pdf See Bartholomew *et al.*, 2005.

FDA, 2003. Guidance for Industry 152 - Evaluating the Safety of Antimicrobial New Animal Drugs with Regard to Their Microbiological Effects on Bacteria of Human Health Concern, October 23, 2003 http://www.fda.gov/cvm/guidance/fguide152.pdf

FDA-CVM, 2004. U S Food and Drug Administration, Center for Veterinary Medicine. Draft Risk Assessment of Streptogramin Resistance in Enterococcus faecium Attributable to the Use of Streptogramins in Animals, November 23, 2004

Fedorka-Cray, PJ. ES Adams, JR Plumblee, LH Dillard, GM Robinson, and ML Headrick. 2001. Changes in Antimicrobial Resistance in *Campylobacter* Isolated from Chicken Carcass Rinses from 1998-2000. Poster Session. http://www.arru.saa.ars.usda.gov/publications/posters/narms/narms_carciccr.pdf

Feldman RA. Confounding factors in observational and intervention studies. Ital J Gastroenterol Hepatol. 1998 Oct;30 Suppl 3:S248-53.

Feller W. *An Introduction to Probability Theory and Its Applications. Volume 1. Third Edition.* John Wiley & Sons. New York. 1968

Fiedler K, Walther E, Freytag P, Nickel S. Inductive reasoning and judgment interference: experiments on Simpson's paradox. Pers Soc Psychol Bull. 2003 Jan;29(1):14-27.

Finch MJ, Blake PA. Foodborne outbreaks of campylobacteriosis: the United States experience, 1980-1982. Am J Epidemiol. 1985 Aug;122(2):262-8.

Follet G. Antibiotic resistance in the EU – Science, politics and policy. *AgriBioForum* 2000. 3(2&3):148-153.

Freedman DA. Graphical models for causation, and the identification problem. Eval Rev. 2004 Aug;28(4):267-93.

Frey L, Fisher D, Tsamardinos I, Aliferis C, Statnikov A. 2003. Identifying Markov Blankets with Decision Tree Induction**.** http://citeseer.nj.nec.com/frey03identifying.html

Friedman CR, Neimann J, Wegener HC, Tauxe RV. Epidemiology of *Campylobacter Jejuni* infections in the United States and Other Industrialized Nations. In Nachamkin I and Blaser MJ, *Campylobacter*, 2nd Ed. ASM Press, Washington, D.C., 2000a. 121-138.

Friedman CR, Hoekstra RM, Samuel M, Marcus R, Bender J, Shiferaw B, Reddy S, Ahuja SD, Helfrick DL, Hardnett F, Carter M, Anderson B, Tauxe RV; Emerging Infections Program FoodNet Working Group. Risk factors for sporadic *Campylobacter* infection in the United States: A case-control study in FoodNet sites. Clin Infect Dis. 2004 Apr 15;38 Suppl 3:S285-96.

Friedman C, Reddy S, Samual M, Marcus R, Bender J, Desai S, Shiferaw B, Helfrick D, Carter M, Anderson B, Hoekstra M, and the EIP Working Group. Risk Factors for Sporadic *Campylobacter* Infections in the United States: A Case-Control Study on FoodNet Sites. 2nd International Conference on Emerging Infectious Diseases. Atlanta, GA, July 2000. http://www.cdc.gov/foodnet/pub/iceid/2000/friedman_c.htm

Friedman N, Goldszmidt, M Learning Bayesian Networks With Local Structure. http://citeseer.ist.psu.edu/friedman96learning.html

Friedman N, Goldszmidt, M. Discretizing Continuous Attributes While Learning Bayesian Networks. 1996b. http://citeseer.ist.psu.edu/friedman96discretizing.html

Frost JA, Gillespie IA, O'Brien SJ. Public health implications of *Campylobacter* outbreaks in England and Wales, 1995-9: epidemiological and microbiological investigations. Epidemiol Infect. 2002 Apr;128(2):111-8.

Furlong, WJ, DH Feeny, GW Torrance, and RD Barr. 2001. The Health Utilities Index system for assessing health-related quality of life in clinical studies. *Annals of Medicine*, 33, 375-384.

FSIS. 1999. "USDA Issues Final Rule on Meat and Poultry Irradiation". Backgrounders, Food Safety and Inspection Service, US Department of Agriculture, Dec 1999.

FSIS. 2000. "Irradiation of Raw Meat and Poultry: Questions and Answers". Consumer Publications, Food Safety and Inspection Service, US Department of Agriculture, May 2000

FSRC, 2003. Food Safety Research Consortium: Ranking the Public Health Impact of Foodborne Hazards, http://www.rff.org/fsrc/riskrankingconference.htm?cf0A52005C=QkFZRVIxXE1PSVJEOmJheWVyMToU5aRGtpBbt/MPPTCqT652

Fugelsang JA, Thompson VA. A dual-process model of belief and evidence interactions in causal reasoning. Mem Cognit. 2003 Jul;31(5):800-15.

Gaudreau C, Gilbert H. Antimicrobial resistance of clinical strains of *Campylobacter* jejuni subsp. jejuni isolated from 1985 to 1997 in Quebec, Canada. Antimicrob Agents Chemother. 1998 Aug;42(8):2106-8.

Gaudreau C, Gilbert H. Antimicrobial resistance of *Campylobacter* jejuni subsp. jejuni strains isolated from humans in 1998 to 2001 in Montreal, Canada. Antimicrob Agents Chemother. 2003 Jun;47(6):2027-9.

Gaudreau C, Michaud S. Cluster of erythromycin- and ciprofloxacin-resistant *Campylobacter* jejuni subsp. jejuni from 1999 to 2001 in men who have sex with men, Quebec, Canada. Clin Infect Dis. 2003 Jul 1;37(1):131-6. Epub 2003 Jun 25.

Gay, K, K Joyce, J Stevenson, F Angulo, T Barrett and the NARMS Working Group. 2002. Quinipristin/dalfopristin-resistant Enterococcus. faecium isolated from human stools, retail chicken, and retail pork: EIP Enterococci Project. International Conference on Emerging Resistant Diseases. Atlanta, GA, March 2002. http://www.cdc.gov/narms/pub/presentations/iceid/gay_k.htm)

Ge B, White DG, McDermott PF, Girard W, Zhao S, Hubert S, Meng J. Antimicrobial-resistant *Campylobacter* species from retail raw meats. Appl Environ Microbiol. 2003 May;69(5):3005-7.

Gefeller O, Land M, Eide GE. Averaging attributable fractions in the multifactorial situation: assumptions and interpretation. J Clin Epidemiol. 1998 May;51(5):437-41.

Gelman A, Van Mechelen I, Verbeke G, Heitjan DF, Meulders M. Multiple imputation for model checking: completed-data plots with missing and latent data. Biometrics. 2005 Mar;61(1):74-85.

George BA, Quarles CL, Fagerberg DJ. Virginiamycin effects on controlling necrotic enteritis infection in chickens. Poult Sci. 1982 Mar;61(3):447-50.

Gillespie IA, O'Brien SJ, Adak GK, Tam CC, Frost JA, Bolton FJ, Tompkins DS; *Campylobacter* Sentinel Surveillance Scheme Collaborators. Point source outbreaks of *Campylobacter* jejuni infection--are they more common than we think and what might cause them? Epidemiol Infect. 2003 Jun;130(3):367-75.

Ginevan, ME. Assessment of the National Antimicrobial Monitoring System (NARMS) and its value in critical decision-making. *International Journal of Infectious Diseases,* December, 2002.

Goldman S. Computational Learning Theory: Lecture Notes. Computer Science Department. Washington University. St. Louis, Missouri. 1991.
http://www.cs.wustl.edu/~sg/CS527_SP01/learning-theory-notes.pdf

Godoy P, Artigues A, Nuin C, Aramburu J, Perez M, Dominguez A, Salleras L. [Outbreak of gastroenteritis caused by *Campylobacter* jejuni transmitted through drinking water] Med Clin (Barc). 2002 Nov 23;119(18):695-8. Spanish. Erratum in: Med Clin (Barc). 2003 Feb 15;120(5):174.

Goossens H. Spread of vancomycin-resistant enterococci: differences between the United States and Europe. Infect Control Hosp Epidemiol. 1998 Aug;19(8):546-51.

Green, PJ. 1995. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. http://citeseer.nj.nec.com/green95reversible.html

Greenland S. Quantifying biases in causal models: classical confounding vs collider-stratification bias. Epidemiology. 2003 May;14(3):300-6.

Greenland S. Modeling and variable selection in epidemiologic analysis. Am J Public Health. 1989 Mar;79(3):340-9.

Greenland S, Brumback B. An overview of relations among causal modeling methods. Int J Epidemiol. 2002 Oct;31(5):1030-7.

Greenland S, Morgenstern H. Confounding in health research. Annu Rev Public Health. 2001;22:189-212.

Grimes DA, Schulz KF. Bias and causal associations in observational research. Lancet. 2002 Jan 19;359(9302):248-52.

Grove WM, Zald DH, Lebow BS, Snitz BE, Nelson C. Clinical versus mechanical prediction: a meta-analysis. Psychol Assess. 2000 Mar;12(1):19-30.

Guatama T, Van Hulle MM. 2003. Surrogate-Based Test For Granger – Causality. http://citeseer.ist.psu.edu/588339.html

Gurm HS, Litaker DG. Framing procedural risks to patients: is 99% safe the same as a risk of 1 in 100? *Acad Med* 2000 Aug;75(8):840-842.

Haas C.N., J.B. Rose, and C.P. Gerba. 1999. *Quantitative Microbial Risk Assessment,* John Wiley & Sons, New York.

Haas CN., On modeling correlated random variables in risk assessment, Risk Anal 1999 Dec;19(6):1205-14

Hagmayer Y, Waldmann MR. How temporal assumptions influence causal judgments. Mem Cognit. 2002 Oct;30(7):1128-37. See also Waldmann MR, Hagmayer Y, Estimating causal strength: the role of structural knowledge and processing effort. Cognition. 2001 Nov;82(1):27-58.

Harris NV, Weiss NS, Nolan CM. The role of poultry and meats in the etiology of *Campylobacter* jejuni/coli enteritis. Am J Public Health. 1986 Apr;76(4):407-11.

Harris N, Thompson D, Martin D and Nolan C. 1986. A survey of *Campylobacter* and other bacterial contaminants of premarket chicken and retail poultry and meats, King County, Washington, *American Journal of Public Health*, 76(4), 401-406.

Hartemink AJ, Gifford DK, Jaakkola TS, Young RA. Using graphical models and genomic expression data to statistically validate models of genetic regulatory networks. Pac Symp Biocomput. 2001;:422-33.

Hayes DJ, Jensen HH, 2003. Lessons from the Danish ban on feed-grade antibiotics http://www.choicesmagazine.org/2003-3/2003-3-01.pdf

Hazen, G, 2003. Multiattribute Structure for QALYS. http://fisher.osu.edu/~butler_267/DAPapers/WP030018.pdf

Hein I, Schneck C, Knogler M, Feierl G, Plesss P, Kofer J, Achmann R, Wagner M. *Campylobacter* jejuni isolated from poultry and humans in Styria, Austria: epidemiology and ciprofloxacin resistance. Epidemiol Infect. 2003 Jun;130(3):377-86.

Helms M, Vastrup P, Gerner-Smidt P, Molbak K.. Short and long term mortality associated with foodborne bacterial gastrointestinal infections: A registry based study. *British Medical Journal* 2003 Feb 15;326(7385):357.

Helms M, Simonsen J, Olsen KEP, Mølbak K. Adverse health events associated with antimicrobial drug resistance in *Campylobacter* species: A registry-based cohort study *J Infectious Diseases* 2005:191 April 1; 1050-1055

Henrion M. Propagation of uncertainty by probabilistic logic sampling in Bayes' networks. In J. F. Lemmer & L. N. Kanal (Eds.), *Uncertainty in Artificial Intelligence 2.* Amsterdam: North Holland, 1988.

Henrion M, Pradhan M, Del Favero B, Huang K, Provan G, O'Rorke P. Why is diagnosis using belief networks insensitive to imprecision in probabilities? 1996. http://citeseer.ist.psu.edu/henrion96why.html

Hershberger E, Oprea SF, Donabedian SM, Perri M, Bozigar P, Bartlett P, Zervos MJ. Epidemiology of antimicrobial resistance in enterococci of animal origin. J Antimicrob Chemother. 2005 Jan;55(1):127-130. Epub 2004 Dec 1.

Heuer OE, Pedersen K, Andersen JS, Madsen M. Vancomycin-resistant enterococci (VRE) in broiler flocks 5 years after the avoparcin ban. Microb Drug Resist. 2002 Summer;8(2):133-8.

Hoffman FO, Hammonds JS. Propagation of uncertainty in risk assessments: the need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability. Risk Anal. 1994 Oct;14(5):707-12.

Holcomb DL, Smith MA, Ware GO, Hung YC, Brackett RE, Doyle MP. Comparison of six dose-response models for use with food-borne pathogens. Risk Anal. 1999 Dec;19(6):1091-100.

Hollander R. [In vitro activity of 23 chemotherapeutic agents against *Campylobacter* jejuni/coli strains isolated from feces] Zentralbl Bakteriol Mikrobiol Hyg [A]. 1983 Dec;256(2):196-201. German.

Ibrahim JG, Chen M-H, Lipsitz SR, Herring AH. Missing-data methods for generalized linear models: A comparative review. Journal of the American Statistical Association 2005 Mar;100(469):332-346.

ICGFI. 1999. "Facts about Food Irradiation". Technical Report, International Consultative Group on Food Irradiation (ICGFI), Vienna, Austria.

Ikram R, Chambers S, Mitchell P, Brieseman MA, Ikam OH. A case control study to determine risk factors for *Campylobacter* infection in Christchurch in the summer of 1992-3. N Z Med J. 1994 Oct 26;107(988):430-2.

Infochoice. 2002. Life expectancy calculator. Infochoice LTD, Sydney, NSW, Australia, (http://www.infochoice.com.au/insurance/health/calculators/lifeexpect.asp).

Iovine NM, Blaser MJ. Antimicrobial resistance in Campylobacter. Emerg Infect Dis. 2004 Jul;10(7):1346.

Iversen A, Kuhn I, Rahman M, Franklin A, Burman LG, Olsson-Liljequist B, Torell E, Mollby R. Evidence for transmission between humans and the environment of a nosocomial strain of *Enterococcus faecium.* Environ Microbiol. 2004 Jan;6(1):55-9.

Iversen A, Kuhn I, Franklin A, Mollby R. High prevalence of vancomycin-resistant enterococci in Swedish sewage. Appl Environ Microbiol. 2002 Jun;68(6):2838-42.

Izat AL, Gardner FA, Denton JH, and Golan FA. 1988. Incidence and level of *Campylobacter* jejuni in broiler processing, *Poultry Science*, 67(11), 1568-72.

Janson S. Coupling And Poisson Approximation. 1994. http://citeseer.ist.psu.edu/janson94coupling.html

Jardine C, Hrudey S, Shortreed J, Craig L, Krewski D, Furgal C, McColl S. Risk management frameworks for human health and environmental risks. J Toxicol Environ Health B Crit Rev. 2003 Nov-Dec;6(6):569-720.

Jaykus LA. The application of quantitative risk assessment to microbial food safety risks. Crit Rev Microbiol. 1996;22(4):279-93.

Jensen LB, Hammerum AM, Bager F, Aarestrup FM. Streptogramin resistance among Enterococcus faecium isolated from production animals in Denmark in 1997. Microb Drug Resist. 2002 Winter;8(4):369-74.

JETACAR. 1999. The use of antibiotics in food-producing animals: antibiotic-resistant bacteria in humans and animals. Report of the Joint Expert Advisory Committee on Antibiotic Resistance (JETACAR). (Commonwealth Department of Health and Aged Care and the Commonwealth Department of Agriculture, Fisheries, and Forestry – Australia) (http://www.health.gov.au/pubs/jetacar.pdf)

Jones FT, Axtell RC, Rives DV, Scheideler FR, Tarver Jr. FR, Walker RL, and Wineland MJ. 1991. A survey of *Campylobacter* jejuni Contamination in Modern Broiler Production and Processing Systems, *Journal of Food Protection*, 54(4), 259-262.

Jones PE, Roelofsma PH. The potential for social contextual and group biases in team decision-making: biases, conditions and psychological mechanisms. Ergonomics. 2000 Aug;43(8):1129-52.

Jones RN, Hare RS, Sabatelli FJ, Ziracin Susceptibility Testing Group. In vitro gram-positive antimicrobial activity of evernimicin (SCH 27899), a novel oligosaccharide, compared with other antimicrobials: a multicentre international trial. *J Antimicrob Chemother* 2001;47:15-25.

Jones, RN, CH Ballow, J Acar, A Fluit, HS Sader, J Turnidge, J Deinhart, and DJ Biedenbach. 1999. World-Wide Evaluation of Quinupristin/Dalfopristin (Q/D; Synercid®) Antimicrobial Activity Directed Against Resistant Gram-Positive Pathogens: Report from the Global SMART Study. Presentation at the 39th ICAAC, September 1999, San Francisco, CA. (http://www.ket-on-line.com/icaac/Posters/Jones.pdf).

Jones, RN, HS Sader, ME Erwin, SA Anderson, and the *Enterococcus* Study Group. 1995. Emerging multiply resistant enterococci among clinical isolates. I. Prevalence data from a 97 medical center surveillance study in the United States. *Diagn. Microbiol. Infect. Dis*, 21, 85–93.

Jordan FT, Forrester CA, Ripley PH, Burch DG. In vitro and in vivo comparisons of valnemulin, tiamulin, tylosin, enrofloxacin, and lincomycin/spectinomycin against Mycoplasma gallisepticum. Avian Dis. 1998 Oct-Dec;42(4):738-45.

Juckett G. Prevention and treatment of traveler's diarrhea. Am Fam Physician. 1999 Jul;60(1):119-24, 135-6.

Kapperud G, Espeland G, Wahl E, Walde A, Herikstad H, Gustavsen S, Tveit I, Natas O, Bevanger L, Digranes A. Factors associated with increased and decreased risk of *Campylobacter* infection: a prospective case-control study in Norway. Am J Epidemiol. 2003 Aug 1;158(3):234-42.

Kang SH, Kodell RL, Chen JJ. Incorporating model uncertainties along with data uncertainties in microbial risk assessment. Regul Toxicol Pharmacol. 2000 Aug;32(1):68-72.

Karmali MA, De Grandis S, Fleming PC.Antimicrobial susceptibility of *Campylobacter* jejuni with special reference to resistance patterns of Canadian isolates. Antimicrob Agents Chemother. 1981 Apr;19(4):593-7.

Kassenborg HD, Smith KE, Vugia DJ, Rabatsky-Ehr T, Bates MR, Carter MA, Dumas NB, Cassidy MP, Marano N, Tauxe RV, Angulo FJ; Emerging Infections Program FoodNet Working Group. Fluoroquinolone-resistant *Campylobacter* infections: eating poultry outside of the home and foreign travel are risk factors. Clin Infect Dis. 2004 Apr 15;38 Suppl 3:S279-84.

Keep Antibiotics Working, 2002. Put Out to Pasture: Strategy to Prolong Antibiotics' Potency. http://www.keepantibioticsworking.com/News/news.cfm?News_ID=207

Keiding N, Budtz-Jorgensen E. The Precautionary Principle and statistical approaches to uncertainty. Int J Occup Med Environ Health. 2004;17(1):147-51.

Kemp, GK, ML Aldrich, M Guerra, and KR Schneider. 2000. “Continuous On-line Processing of Fecal Contaminated Poultry Carcasses”. Presented at IAFP, Atlanta, GA, August, 2000.

Kirk M, Waddell R, Dalton C, Creaser A, Rose N. A prolonged outbreak of Campylobacter infection at a training facility. Commun Dis Intell. 1997 Mar 6;21(5):57-61.

Klare I, Badstubner D, Konstabel C, Bohme G, Claus H, Witte W. Decreased incidence of VanA-type vancomycin-resistant enterococci isolated from poultry meat and from fecal samples of humans in the community after discontinuation of avoparcin usage in animal husbandry. Microb Drug Resist. 1999 Spring;5(1):45-52.

Klein G, Pack A, Reuter G. Antibiotic resistance patterns of enterococci and occurrence of vancomycin-resistant enterococci in raw minced beef and pork in Germany. Appl Environ Microbiol. 1998 May;64(5):1825-30.

Knothe H. Medical implications of macrolide resistance and its relationship to the use of tylosin in animal feeds. Infection. 1977;5(3):137-9.

Koenraad PM, Jacobs-Reitsma WF, Van der Laan T, Beumer RR, Rombouts FM. Antibiotic susceptibility of *Campylobacter* isolates from sewage and poultry abattoir drain water. Epidemiol Infect. 1995 Dec;115(3):475-83.

Lange K. *Applied Probability.* Springer. New York. 2003

Lautenbach, E, WB Bilker, and PJ Brennan. 1999. Enterococcal bacteremia: risk factors for vancomycin resistance and predictors of mortality. *Infect Control Hosp Epidemiol*, 20(5), 318-23. (http://www.slackinc.com/general/iche/ichehome.htm)

Lawton RM, Fridkin SK, Gaynes RP, McGowan JE Jr. Practices to improve antimicrobial use at 47 US hospitals: the status of the 1997 SHEA/IDSA position paper recommendations. Society for Healthcare Epidemiology of America/Infectious Diseases Society of America. Infect Control Hosp Epidemiol. 2000 Apr;21(4):256-9. (http://www.slackinc.com/general/iche/ichehome.htm)

Lemck R, Bulte M. Occurrence of the vancomycin-resistant genes vanA, vanB, vancl, vanC2 and vanC3 in Enterococcus strains isolated from poultry and pork. Int J Food Microbiol. 2000 Sep 25;60(2-3):185-94.

Lemon SC, Roy J, Clark MA, Friedmann PD, Rakowski W. Classification and regression tree analysis in public health: methodological review and comparison with logistic regression. Ann Behav Med. 2003 Dec;26(3):172-81.

Lighton LL, Kaczmarski EB, Jones DM. A study of risk factors for *Campylobacter* infection in late spring. Public Health. 1991 May;105(3):199-203.

Linden, PK. 2002. Treatment Options for Vancomycin- Resistant Enterococcal Infections. Drugs, 62(3), 425-441.

Linden PK, AW Pasculle, D McDevitt, and DJ Kramer. 1997. Effect of quinupristin/dalfopristin on the outcome of vancomycin-resistant Enterococcus faecium bacteraemia: comparison with a control cohort. *J Antimicrob Chemother*, 39 Suppl A, 145-51. (http://www.jac.oupjournals.org)

Linkov I, Burmistrov D. Model uncertainty and choices made by modelers: lessons learned from the International Atomic Energy Agency model intercomparisons. Risk Anal. 2003 Dec;23(6):1297-308.

Lillard HS. 1990. The impact of commercial processing procedures on bacterial contamination and cross-contamination of broiler carcasses, *Journal of Food Protection*, 53(3), 202-204.

Loh, 2002. Regression trees with unbiased variable selection and interaction. Statistica Sinica **12**(2002), 361-386. www.stat.sinica.edu.tw/statistica/password.asp?vol=12&num=2&art=1

Lovland, A and M Kaldhusdal. 2001. Severely impaired production performance in broiler flocks with high incidence of *Clostridium perfringens*-associated hepatitis. *Avian Pathology*, 30, 73-81.
http://www.poultry-health.com/fora/inthelth/kaldhusdal02.htm

Luber P, Wagner J, Hahn H, Bartelt E. Antimicrobial resistance in *Campylobacter* jejuni and *Campylobacter* coli strains isolated in 1991 and 2001-2002 from poultry

and humans in Berlin, Germany. Antimicrob Agents Chemother. 2003 Dec;47(12):3825-30.

Luce, RD and P Suppes. 2001. Representational Measurement Theory. http://media.wiley.com/product_data/excerpt/87/04713788/0471378887.pdf.

Maclure M. Multivariate refutation of aetiological hypotheses in non-experimental epidemiology. Int J Epidemiol. 1990 Dec;19(4):782-7.

Maclure M.Taxonomic axes of epidemiologic study designs: a refutationist perspective. J Clin Epidemiol. 1991;44(10):1045-53.

MacGregor DG, Slovic P, Malmfors T. How exposed is exposed enough? Lay inferences about chemical exposure. *Risk Anal* 1999 Aug;19(4):649-659.

Madsen, M and K Pederson. 2000. Research activities on Clostridiosis and Coccidiosis in broilers in Denmark, Proceedings of the *AFAC Workshop on AClostridiosis and Coccidiosis in Broilers*, Uppsala, Denmark, Sept. 14-15, 2000. http://www-afac.slu.se/Madsen_Pedersen.pdf

Manson JM, Smith JM, Cook GM. Persistence of vancomycin-resistant enterococci in New Zealand broilers after discontinuation of avoparcin use. Appl Environ Microbiol. 2004 Oct;70(10):5764-8.

Marano, N, D. Vugia,, T. Fiorentino, S. Segler, M. Carter, H. Kassenborg, K. Smith, S. Sansky, K. Hollinger, F. Angulo, and the EIP FoodNet Working Group. 2000, "Fluoroquinolone-Resistant *Campylobacter* Causes Longer Duration of Diarrhea than Fluoroquinolone-susceptible *Campylobacter* Strains in FoodNet Sites". Presented at the International Conference on Emerging Infectious Diseases, 2000 July, Atlanta, GA.
http://www.cdc.gov/narms/pub/presentations/2000/marano_n_3.htm

Mark DH. Interpreting the term selection bias in medical research. Fam Med. 1997 Feb;29(2):132-6.

Marks HM, Coleman ME, Lin CT, Roberts T. Topics in microbial risk assessment: dynamic flow tree process. Risk Analysis 1998 Jun;18(3):309-28.

Mather FJ, White LE, Langlois EC, Shorter CF, Swalm CM, Shaffer JG, Hartley WR. Statistical methods for linking health, exposure, and hazards. Environ Health Perspect. 2004 Oct;112(14):1440-5.
http://ehp.niehs.nih.gov/members/2004/7145/7145.html

McCloy R, Byrne RM. Counterfactual thinking about controllable events. Mem Cognit. 2000 Sep;28(6):1071-8.

McDermott PF, Bodeis SM, English LL, White DG, Walker RD, Zhao S, Simjee S, Wagner DD. Ciprofloxacin resistance in *Campylobacter* jejuni evolves rapidly in chickens treated with fluoroquinolones. J Infect Dis. 2002 Mar 15;185(6):837-40. Epub 2002 Feb 08.

McDonald LC, Rossiter S, Mackinson C, Wang YY, Johnson S, Sullivan M, Sokolow R, DeBess E, Gilbert L, Benson JA, Hill B, Angulo FJ 2001. Quinupristin-dalfopristin-resistant Enterococcus faecium on chicken and in human stool specimens. N Engl J Med. Oct 18;345(16):1155-60.

McNab B, Alves D. Risk Assessment Frameworks - Animal Health Risk Assessment. Ontario Ministry of Agriculture and Food (OMAF). 2003. www.gov.on.ca/OMAFRA/english/research/risk/frameworks/as3c.html#detailed www.gov.on.ca/OMAFRA/english/research/risk/frameworks/as2c.html#1.0%20Risk%20Assessment1841

Mead PS, Slutsker L, Dietz V, McCaig LF, Bresee JS, Shapiro C, Griffin PM, Tauxe RV.. 1999. Food-related illness and death in the United States. *Emerging Infectious Diseases*, 5(5), Sep-Oct, 607-625 http://www.cdc.gov/ncidod/EID/vol5no5/mead.htm

Mead GC, Hudson WR, Hinton MH. Effect of changes in processing to improve hygiene control on contamination of poultry carcasses with campylobacter. Epidemiol Infect. 1995 Dec;115(3):495-500.

Medema GJ, Teunis PF, Havelaar AH, Haas CN. Assessment of the dose-response relationship of *Campylobacter* jejuni. Int J Food Microbiol. 1996 Jun;30(1-2):101-11.

Michaud S, Menard S, Arbeit RD. Campylobacteriosis, Eastern Townships, Quebec. Emerg Infect Dis. 2004 Oct;10(10):1844-7.

Michaud S, Menard S, Ahei RD. Risk Factors for *Campylobacter* Infections in Estrie, Quebec: A Case-Control Study. 11th International Workshop on Campylobacter, Helicobacter and Related Organisms. International Journal of Medical Microbiology. http://193.196.199.11/chro2001/

Michel J, Rogol M, Dickman D. Susceptibility of clinical isolates of *Campylobacter* jejuni to sixteen antimicrobial agents. Antimicrob Agents Chemother. 1983 May;23(5):796-7.

Miyamoto JM. Quality-Adjusted Life Years (QALY) Utility Models under Expected Utility and Rank Dependent Utility Assumptions. J Math Psychol. 1999 Jun;43(2):201-237.

MLA. 2003. Meat and Livestock Australia. http://www.mla.com.au/content.cfm

Moellering RC, Linden PK, Reinhardt J, Blumberg EA, Bompart F, Talbot GH. The efficacy and safety of quinupristin/dalfopristin for the treatment of infections caused by vancomycin-resistant Enterococcus faecium. Synercid Emergency-Use Study Group. J Antimicrob Chemother. 1999 Aug;44(2):251-61.

Molbak K. Spread of resistant bacteria and resistance genes from animals to humans -- the public health consequences. J Vet Med B Infect Dis Vet Public Health. 2004 Oct-Nov;51(8-9):364-9.

Moore JE, O'Riordan L, Wareing DR, Doyle R, Lanser J, Stanley T, Matsuda M, Matsui T, Murphy PG. Phenotypic and genotypic relationship between *Campylobacter* spp isolated from humans and chickens in Northern Ireland--a comparison of three phenotyping and two genotyping schemes. Int J Hyg Environ Health. 2003 Jun;206(3):211-6

Morris MW, Moore PC, Sim DL. Choosing remedies after accidents: counterfactual thoughts and the focus on fixing "human error". Psychon Bull Rev. 1999 Dec;6(4):579-85.

Murray, BE. 2000. Vancomycin-resistant enterococcal infections. *The New England Journal of Medicine*, Mar. 9, 342(10), pp 710-721. (http://content.nejm.org/)

Mulder, RW. 1997. “Safe Poultry Meat Production in the Next Century”. *Acta Vet Hung* 1997;45(3):307-15.

Musgrove MT, Cox NA, Berrang ME, Harrison MA. Comparison of weep and carcass rinses for recovery of *Campylobacter* from retail broiler carcasses. J Food Prot. 2003 Sep;66(9):1720-3.

Mutnick, AH, V Enne, and RN Jones. Linezolid resistance since 2001: SENTRY Antimicrobial Surveillance Program. *Ann Pharmacother*, 37(6), 909-11.

Nadeau E, Messier S, Quessy S.. Prevalence and comparison of genetic profiles of *Campylobacter* strains isolated from poultry and sporadic cases of campylobacteriosis in humans. J Food Prot. 2002 Jan;65(1):73-8.

NCHS, 2001. United States Life Tables, 1998. CDC, National Center for Health Statistics (NCHS), National Vital Statistics System, National Vital Statistics Reports, 48(18), October, 2001. (http://www.cdc.gov/nchs/data/nvsr/nvsr48/nvs48_18.pdf)

Neapolitan RE. The principle of interval constraints: a generalization of the symmetric Dirichlet distribution. Math Biosci. 1991 Feb;103(1):33-44.

Nelson JM, Smith KE, Vugia DJ, Rabatsky-Ehr T, Segler SD, Kassenborg HD, Zansky SM, Joyce K, Marano N, Hoekstra RM, Angulo FJ. Prolonged Diarrhea

Due to Ciprofloxacin-Resistant *Campylobacter* Infection. J Infect Dis. 2004 Sep 15;190(6):1150-7. Epub 2004 Aug 03.

Ng KL, Hamby DM.. Fundamentals for establishing a risk communication program. *Health Phys* 1997 Sep;73(3):473-482.

NNIS. 2001. National Nosocomial Infections Surveillance (NNIS) System Report, Data Summary from January 1992-June 2001, Issued August 2001, *American Journal of Infection Control*, 29. 404-421. (http://www.apic.org/ajic/) or for raw report:(http://www.cdc.gov/ncidod/hip/NNIS/members/2001NNIS_report.pdf)

NNIS. 2000. SemiAnnual Report, Aggregated Data from the National Nosocomial Infections Surveillance (NNIS) System, corrected version issued 3/29/2000. (http://www.cdc.gov/ncidod/hip/NNIS/dec99sar.pdf)

Nolan CM, Johnson KE, Coyle MB, Faler K. *Campylobacter* jejuni enteritis: efficacy of antimicrobial and antimotility drugs. Am J Gastroenterol. 1983 Oct;78(10):621-6.

Oscar TP. A quantitative risk assessment model for *Salmonella* and whole chickens. Int J Food Microbiol. 2004 Jun 1;93(2):231-47.

Ottenbacher KJ. Quantitative evaluation of multiplicity in epidemiology and public health research. Am J Epidemiol. 1998 Apr 1;147(7):615-9.

Owens DK, Shachter RD, Nease RF Jr. Representation and analysis of medical decision problems with influence diagrams. Med Decis Making. 1997 Jul-Sep;17(3):241-62.

Padiglione, A. 2001. Vancomycin resistant Enterococcus in hospital patients. http://www.med.monash.edu/epidemiology/units/infdis/enterococcus.html and personal correspondence from A. Padiglione.

Parsons DJ, Orton TG, D'Souza J, Moore A, Jones R, Dodd CE. A comparison of three modeling approaches for quantitative risk assessment using the case study of *Salmonella* spp. in poultry meat. Int J Food Microbiol. 2005 Jan 15;98(1):35-51.

Patton DE. The ABCs of risk assessment. EPA JOURNAL 1993 19(1):10-15 http://www.bethel.edu/~kisrob/hon301k/readings/risk/RiskEPA/riskepa1.html

Pearl J. Causal Inference in the Health Sciences: A Conceptual Introduction. Health Services Outcomes Research Methodology 2:189-220, 2002. http://citeseer.ist.psu.edu/599949.html

Pebody RG, Ryan MJ, Wall PG. Outbreaks of campylobacter infection: rare events for a common pathogen. Commun Dis Rep CDR Rev. 1997 Mar 7;7(3):R33-7.

Pezzotti G, Serafin A, Luzzi I, Mioni R, Milan M, Perin R. Occurrence and resistance to antibiotics of *Campylobacter* jejuni and *Campylobacter* coli in animals and meat in northeastern Italy. Int J Food Microbiol. 2003 May 15;82(3):281-7.

Pharmacia & Upjohn, 2000. Linezolid-A New Therapeutic Option for Managing Known or Suspected Gram-Positive Infections. Presentation to the Speakers Club of the Communications Center for Infectious Diseases, Alan Tice (ed.), http://www.idlinks.com/spkrclub/slides_antibiotics.htm.

Phillips I, Casewell M, Cox T, De Groot B, Friis C, Jones R, Nightingale C, Preston R, Waddell J. Does the use of antibiotics in food animals pose a risk to human health? A critical review of published data. J Antimicrob Chemother. 2004 Jan;53(1):28-52.

Phillips CV, Goodman KJ. The missed lessons of Sir Austin Bradford Hill. Epidemiol Perspect Innov. 2004 Oct 04;1(1):3.

Pichler H, Diridl G, Wolf D. Ciprofloxacin in the treatment of acute bacterial diarrhea: a double blind study. Eur J Clin Microbiol. 1986 Apr;5(2):241-3.

Piddock LJV. 1999. Implications for Human Health, *The Journal of Antimicrobial Chemotherapy, Suppl A,* 44:17.

Plous S. *The Psychology of Judgment and Decision Making.* 1993.

Potter RC, Kaneene JB, Hall WN. Risk factors for sporadic *Campylobacter* jejuni infections in rural Michigan: a prospective case-control study. Am J Public Health. 2003 Dec; 93(12): 2118-23.

RAC (Risk Assessment Consortium), Data Gaps Working Group. Data Gaps for Selected Microbial Risk Assessments. December 02, 2004. http://www.foodrisk.org/RAC_data_gaps.htm

Rai SN, Krewski D. Uncertainty and variability analysis in multiplicative risk models. Risk Anal. 1998 Feb;18(1):37-45.

Rapoport J, Teres D, Zhao Y, Lemeshow S.. Length of stay data as a guide to hospital economic performance for ICU patients. Med Care. 2003 Mar;41(3):386-97.

Rautelin H, Vierikko A, Hanninen ML, Vaara M. Antimicrobial susceptibilities of *Campylobacter* strains isolated from Finnish subjects infected domestically or from those infected abroad. Antimicrob Agents Chemother. 2003 Jan;47(1):102-5.

Rautelin H, Renkonen OV, Kosunen TU. Emergence of fluoroquinolone resistance in *Campylobacter* jejuni and *Campylobacter* coli in subjects from Finland. Antimicrob Agents Chemother. 1991 Oct;35(10):2065-9.

Reina J, Ros MJ, Serra A. Susceptibilities to 10 antimicrobial agents of 1,220 *Campylobacter* strains isolated from 1987 to 1993 from feces of pediatric patients. Antimicrob Agents Chemother. 1994 Dec;38(12):2917-20.

Rice, LB. 2001. Emergence of Vancomycin Resistant Enterococci, *Emerging Infectious Diseases*, 7(2), 183-187.

Richardson S, Green PJ, 1997. On Bayesian analysis of mixtures with an unknown number of components. http://citeseer.ist.psu.edu/richardson97bayesian.html

Robinson, DA. 1981. "Infective Dose of *Campylobacter* jejuni in Milk", *British Medical Journal*, 282, 1584.

Roberts MC, Sutcliffe J, Courvalin P, Jensen LB, Rood J, Seppala H. Nomenclature for macrolide and macrolide-lincosamide-streptogramin B resistance determinants. Antimicrob Agents Chemother. 1999 Dec;43(12):2823-30.

Robins-Browne RM, Mackenjee MK, Bodasing MN, Coovadia HM. Treatment of Campylobacter-associated enteritis with erythromycin. Am J Dis Child. 1983 Mar;137(3):282-5.

Rodrigues LC, Cowden JM, Wheeler JG, Sethi D, Wall PG, Cumberland P, Tompkins DS, Hudson MJ, Roberts JA, Roderick PJ. The study of infectious intestinal disease in England: risk factors for cases of infectious intestinal disease with *Campylobacter* jejuni infection. Epidemiol Infect. 2001 Oct;127(2):185-93.

Roels TH, Wickus B, Bostrom HH, Kazmierczak JJ, Nicholson MA, Kurzynski TA, Davis JP. A foodborne outbreak of *Campylobacter* jejuni (O:33) infection associated with tuna salad: a rare strain in an unusual vehicle. Epidemiol Infect. 1998 Oct;121(2):281-7.

Romano JP, Wolf M. Exact and approximate stepdown methods for multiple hypothesis testing. Journal of the American Statistical Association 2005 Mar;100(469):94-108.

Rosef O, Kapperud G, Lauwers S, Gondrosen B. Serotyping of *Campylobacter* jejuni, *Campylobacter* coli, and *Campylobacter* laridis from domestic and wild animals. Appl Environ Microbiol. 1985 Jun;49(6):1507-10.

Rosenfield JA, Arnold GJ, Davey GR, Archer RS, Woods WH. Serotyping of *Campylobacter* jejuni from an outbreak of enteritis implicating chicken. J Infect. 1985 Sep;11(2):159-65.

Rosenquist H, Nielsen NL, Sommer HM, Norrung B, Christensen BB. Quantitative risk assessment of human campylobacteriosis associated with thermophilic *Campylobacter* species in chickens. Int J Food Microbiol. 2003 May 25;83(1):87-103.

Ross, SM. 1996. *Stochastic Processes, Second Edition.* Wiley, New York.

Ross T, McMeekin TA. Modeling microbial growth within food safety risk assessments. Risk Anal. 2003 Feb;23(1):179-97.

Rossiter S, Joyce K, Ray M, Benson J, Mackinson C, Gregg C, Sullivan M, Vought K, Leano F, Besser J, Marano N, and Angulo F. 2000. High Prevalence of antimicrobial-resistant, including fluoroquinolone-resistant *Campylobacter* on chicken in US grocery stores, American Society for Microbiology 100th General Meeting, 2000 May, Los Angeles CA. http://www.cdc.gov/ncidod/dbmd/narms/pub/presentations/2000/rossiter_s_3.htm

Russell SM. The effect of airsacculitis on bird weights, uniformity, fecal contamination, processing errors, and populations of *Campylobacter* spp. and Escherichia coli. Poult Sci. 2003 Aug;82(8):1326-31.

Russell, SM. "Poultry Health Impact on Food Safety", Proceedings of the Poultry Health and Processing Conference, Delmarva Poultry Processing, Ocean City, MD, Oct 11, 2002.

Sader HS, Streit JM, Fritsche TR, Jones RN. Antimicrobial activity of daptomycin against multidrug-resistant Gram-positive strains collected worldwide. Diagn Microbiol Infect Dis. 2004 Nov;50(3):201-4.

Saltelli A, Tarantola S, Campolongo F, Ratto M. *Sensitivity Analysis in Practice.* Wiley, 2004.

Sanchez R, Fernandez-Baca V, Diaz MD, Munoz P, Rodriguez-Creixems M, Bouza E. Evolution of susceptibilities of *Campylobacter* spp. to quinolones and macrolides. Antimicrob Agents Chemother. 1994 Sep;38(9):1879-82.

Sandman P. Responding to Community Outrage: Strategies for Effective Risk Communication. American Industrial Hygiene Association. 1993.

Savitz DA, Greenland S, Stolley PD, Kelsey JL. Scientific standards of criticism: a reaction to "Scientific standards in epidemiologic studies of the menace of daily life," by A.R. Feinstein. Epidemiology. 1990 Jan;1(1):78-83.

Schafer JL. *Analysis of Incomplete Multivariate Data.* New York: Chapman and Hall;1997

Schouls LM, Reulen S, Duim B, Wagenaar JA, Willems RJ, Dingle KE, Colles FM, Van Embden JD. Comparative genotyping of *Campylobacter* jejuni by amplified

fragment length polymorphism, multilocus sequence typing, and short repeat sequencing: strain diversity, host range, and recombination. J Clin Microbiol. 2003 Jan;41(1):15-26.

Schwaber MJ, De-Medina T, Carmeli Y. Epidemiological interpretation of antibiotic resistance studies - what are we missing? Nat Rev Microbiol. 2004 Dec;2(12):979-83.

Schwartz JA, Chapman GB. Are more options always better? The attraction effect in physicians' decisions about medications. Med Decis Making. 1999 Jul Sep; 19(3):315-23.

Sebastiani, P. 2001. Bayesian Data Analysis. Draft Technical Paper, Dept. of Mathematics and Statistics, University of Massachusetts http://www.math.umass.edu/~sebas

Shipley, B. (2000). Cause and Correlation in Biology. A User's Guide to Path Analysis, Structural Equations and Causal Inference. Cambridge University Press. http://callisto.si.usherb.ca:8080/bshipley/my%20book.htm

Shnayerson M, Plotkin MJ, 2002. *The Killers Within: The Deadly Rise of Drug-Resistant Bacteria.* Little, Brown and Company. Boston.

Sielken RL Jr, Bretzlaff RS, Stevenson DE. Challenges to default assumptions stimulate comprehensive realism as a new tier in quantitative cancer risk assessment. *Regul Toxicol Pharmacol.* 1995 Apr;21(2):270-80.

Simjee S, White DG, Wagner DD, Meng J, Qaiyumi S, Zhao S, McDermott PF. Identification of vat(E) in Enterococcus faecalis isolates from retail poultry and its transferability to Enterococcus faecium. Antimicrob Agents Chemother. 2002 Dec;46(12):3823-8.

Sjogren E, Kaijser B, Werner M. Antimicrobial susceptibilities of *Campylobacter* jejuni and *Campylobacter* coli isolated in Sweden: a 10-year follow-up report. Antimicrob Agents Chemother. 1992 Dec;36(12):2847-9.

Sjogren E, Lindblom GB, Kaijser B. Norfloxacin resistance in *Campylobacter* jejuni and *Campylobacter* coli isolates from Swedish patients. J Antimicrob Chemother. 1997 Aug;40(2):257-61.

Skirrow MB. *Campylobacter* enteritis: a "new" disease. Br Med J. 1977 Jul 2;2(6078):9-11.

Skirrow MB. Epidemiology of *Campylobacter* enteritis. Int J Food Microbiol. 1991 Jan;12(1):9-16.

Smith KE, Besser JM, Hedberg CW, Leano FT, Bender JB, Wicklund JH, Johnson BP, Moore KA, Osterholm MT.. Quinolone-Resistant *Campylobacter* jejuni

Infections in Minnesota, 1992-1998, *New England Journal of Medicine*, vol. 340, No. 20, May 20, 1525-1582.

Smith DL, Johnson JA, Harris AD, Furuno JP, Perencevich EN, Morris JG Jr. Assessing risks for a pre-emergent pathogen: virginiamycin use and the emergence of streptogramin resistance in *Enterococcus faecium.* Lancet Infect Dis. 2003 Apr;3(4):241-9. Details are at: http://medschool.umaryland.edu/Epidemiology/dsmith/sref.htm

Smith DL, Harris AD, Johnson JA, Silbergeld EK, Morris JG Jr. Animal antibiotic use has an early but important impact on the emergence of antibiotic resistance in human commensal bacteria. *Proceedings of the National Academy of Sciences.* Proc Natl Acad Sci U S A. 2002 Apr 30;99(9):6434-9. http://www.pnas.org/cgi/content/full/082188899v1

SNJ. 2000. Epidemiology Surveillance System 1998 Report. New Jersey Department of Health and Senior Services Division of Epidemiology, Environmental and Occupational Health, October 2000. (http://www.state.nj.us/health/cd/ess1998/)

Sonis J. A closer look at confounding. Fam Med. 1998 Sep;30(8):584-8.

Sopwith W, Ashton M, Frost JA, Tocque K, O'Brien S, Regan M, Syed Q. Enhanced surveillance of *Campylobacter* infection in the North West of England 1997-1999. J Infect. 2003 Jan;46(1):35-45.

Sorvillo FJ, Lieb LE, Waterman SH. Incidence of campylobacteriosis among patients with AIDS in Los Angeles County. J Acquir Immune Defic Syndr. 1991;4(6):598-602.

Stalder H, Isler R, Stutz W, Salfinger M, Lauwers S, Vischer W. [Contribution to the epidemiology of *Campylobacter* jejuni. From asymptomatic excretion by a cow in the cowshed to overt disease in over 500 persons] Schweiz Med Wochenschr. 1983 Feb 19;113(7):245-9.

Steffen W, Mochmann H, Kontny I, Richter U, Werner U, el Naeem O. [A food-borne infection caused by *Campylobacter* jejuni serotype Lauwers 19] Zentralbl Bakteriol Mikrobiol Hyg [B]. 1986 Dec;183(1):28-35. German.

Stern NJ, Robach MC. Enumeration of *Campylobacter* spp. in broiler feces and in corresponding processed carcasses. J Food Prot. 2003 Sep;66(9):1557-63.

Stern NJ, Clavero MR, Bailey JS, Cox NA, Robach MC.. 1995a. *Campylobacter* spp. in broilers on the farm and after transport. Poult Sci. 1995 Jun;74(6):937-41..

Stern, NJ. 1995b. Influence of season and refrigerated storage on *Campylobacter* spp. contamination of broiler carcasses, *Journal of Applied Poultry Research*, Vol 4, 235-238.

Studahl A, Andersson Y. Risk factors for indigenous *Campylobacter* infection: a Swedish case-control study. Epidemiol Infect. 2000 Oct;125(2):269-75.

Stewart N, Chater N, Stott HP, Reimers S. Prospect relativity: how choice options influence decision under risk. J Exp Psychol Gen. 2003 Mar;132(1):23-46.

Surgeon General, 2004. The Health Consequences of Smoking: A Report of the Surgeon General

Svedhem A, Kaijser B, Sjogren E, 1981. Antimicrobial susceptibility of *Campylobacter* jejuni isolated from humans with diarrhea and from healthy chickens. Journal of Antimicrobial Chemotherapy, 1981; 7:301-305.

Swanson NR, Ozyildirim A, Pisu M. 2001. A comparison of alternative causality and predictive accuracy tests in the presence of integrated and co-integrated economic variables. http://citeseer.nj.nec.com/swanson01comparison.html

Tammemagi CM, Neslund-Dudas C, Simoff M, Kvale P. Impact of comorbidity on lung cancer survival. Int J Cancer. 2003 Mar 1;103(6):792-802.

Tangen JM, Allan LG. Cue interaction and judgments of causality: contributions of causal and associative processes. Mem Cognit. 2004 Jan;32(1):107-24.

Teuber M, Schwarz F, Perreten V. Molecular structure and evolution of the conjugative multiresistance plasmid pRE25 of Enterococcus faecalis isolated from a raw-fermented sausage. Int J Food Microbiol. 2003 Dec 1;88(2-3):325-9.

Teunis PF, Nagelkerke NJ, Haas CN. 1999. Dose response models for infectious gastroenteritis. Risk Anal. 1999 Dec;19(6):1251-60.

Teunis PF, Havelaar AH.. The Beta Poisson dose-response model is not a single-hit model. Risk Anal. 2000 Aug;20(4):513-20.

Thal, LA, S Donabedian, B Robinson-Dunn, JW Chow, L Dembry, DB Clewell, D Alshab, and MJ Zervos. 1998. Molecular Analysis of Glycopeptide-Resistant *Enterococcus faecium* Isolates Collected from Michigan Hospitals over a 6-Year Period, *Journal of Clinical Microbiology*, 36(11), 3303-3308. (http://jcm.asm.org/)

Thompson KM, Burmaster DE, Crouch EA. Monte Carlo techniques for quantitative uncertainty analysis in public health risk assessments. Risk Anal. 1992 Mar;12(1): 53-63.

Tollefson L. Factual errors in review article. J Antimicrob Chemother. 2004 Jul;54(1):271; author reply 276-8.

Tollefson L, Karp BE. Human health impact from antimicrobial use in food animals. Med Mal Infect. 2004 Nov;34(11):514-21.

Tsamardinos I, Aliferis C, Statnikov A. 2003. Time and Sample Efficient Discovery of Markov Blankets and Direct Causal Relations. http://citeseer.nj.nec.com/tsamardinos03time.html

Turnidge, J. 2001. Antimicrobial resistance surveillance – what happens if we do nothing. Presentation at the Antibiotic Resistance Surveillance Workshop, Melbourne, Australia, 4 May 2001. (http://www.health.gov.au/pubhlth/strateg/jetacar/pdf/turnidge.pdf)

Turnidge J. and Bell J., 2002. Surveillance of resistance to Synercid® (Quinupristin/Dalfopristin) in the Western Pacific, 1998. Poster, Table 5.

US Census Bureau, http://www.census.gov/main/www/popclock.html

USDA. 1988. The US Broiler Industry, USDA Economic Research Service, http://usda.mannlib.cornell.edu/usda/

USDA. 1996. "Bacterial Foodborne Disease: Medical Costs and Productivity Losses". Food and Consumer Economics Division, Economic Research Service, U.S. Department of Agriculture. Agricultural Economic Report No. 741.

USDA. 2000. Poultry Slaughter, Figures released 6/2/2000, USDA National Agricultural Statistics Service, http://usda.mannlib.cornell.edu/reports/nassr/poultry/ppy-bb/2000

USDA-FDA, 2004. USDA/FDA HACCP Training Programs and Resources Database http://www.nal.usda.gov/fnic/foodborne/haccp/index.shtml

VanDer Kop. M, 2002. *Campylobacter jejuni* in chickens at slaughter and retail. http://www.poultryworkshop.com/workshops/2002/2002_abstracts/campylobacter_jejuni.html

van Gerwen SJC, te Giffel MC, van 't Riet K, Beumer RR, Zwietering, MH. Stepwise quantitative risk assessment as a tool for characterization of microbiological food safety. *Journal of Applied Microbiology* 2000 **88** (6), 938-951. doi: 10.1046/j.1365-2672.2000.01059.x http://www.blackwell-synergy.com/links/doi/10.1046/j.1365-2672.2000.01059.x/full/

van Gerwen SJ, Zwietering MH. Growth and inactivation models to be used in quantitative risk assessments. J Food Prot. 1998 Nov;61(11):1541-9.

Vanhoof R, Gordts B, Dierickx R, Coignau H, Butzler JP. Bacteriostatic and bactericidal activities of 24 antimicrobial agents against *Campylobacter* fetus subsp. jejuni. Antimicrob Agents Chemother. 1980 Jul; 18(1): 118-21.

Vellinga A, Van Loock F. The dioxin crisis as experiment to determine poultry-related *Campylobacter* enteritis. Emerg Infect Dis. 2002 Jan;8(1):19-22.

Viallefont V, Raftery AE, Richardson S. Variable selection and Bayesian model averaging in case-control studies. Stat Med. 2001 Nov 15;20(21):3215-30.

Vissiennon T, Kroger H, Kohler T, Kliche R. Effect of avilamycin, tylosin and ionophore anticoccidials on Clostridium perfringens enterotoxaemia in chickens. Berl Munch Tierarztl Wochenschr. 2000 Jan;113(1):9-13.

Veterinary Laboratories Agency. Surveillance Report Avian 2004. Quarterly Report 7(4):9-10 http://www.defra.gov.uk/corporate/vla/science/documents/science-end-survrep-qtlya404.pdf

Vose DJ. The application of quantitative risk assessment to microbial food safety. J Food Prot. 1998 May;61(5):640-8.

Vose, DJ. *Risk Analysis: A Quantitative Guide,* 2nd Ed. John Wiley & Sons. 2000.

Vose DJ, Acar J, Anthony F, Franklin A, Gupta R, Nicholls T, Tamura Y, Thompson S, Threlfall EJ, van Vuuren M, White DG, Wegener HC, Costarrica ML; Office International des Epizooties Ad hoc Group. Antimicrobial resistance: risk analysis methodology for the potential impact on public health of antimicrobial resistant bacteria of animal origin. Rev Sci Tech. 2001 Dec;20(3):811-27.

Walz SE, Baqar S, Beecham HJ, Echeverria P, Lebron C, McCarthy M, Kuschner R, Bowling S, Bourgeois AL, Scott DA. Pre-exposure anti-*Campylobacter* jejuni immunoglobulin a levels associated with reduced risk of *Campylobacter* diarrhea in adults traveling to Thailand. Am J Trop Med Hyg. 2001 Nov;65(5):652-6.

Wang D, Zhang W, Bakhai A. Comparison of Bayesian model averaging and stepwise methods for model selection in logistic regression. Stat Med. 2004 Nov 30;23(22):3451-67.

Webb, M, LW Riley, and RB Roberts. 2001. Cost of hospitalization for and risk factors associated with vancomycin resistant Enterococcus faecium infection and colonization. *Clinical Infectious Diseases*, 33, 4, 445-452. (http://www.journals.uchicago.edu/CID/journal/home.html)

Wedderkopp A, Nielsen EM, Pedersen K. Distribution of *Campylobacter* jejuni Penner serotypes in broiler flocks 1998-2000 in a small Danish community with special reference to serotype 4-complex. Epidemiol Infect. 2003 Oct;131(2):915-21.

Weed DL. Interpreting epidemiological evidence: how meta-analysis and causal inference methods are related. Int J Epidemiol. 2000 Jun;29(3):387-90.

Wegener HC, Aarestrup FM, Jensen LB, Hammerum AM, Bager F. Use of antimicrobial growth promoters in food animals and Enterococcus faecium resistance to therapeutic antimicrobial drugs in Europe. Emerg Infect Dis. 1999 May-Jun;5(3):329-35. Erratum in: Emerg Infect Dis 1999 Nov-Dec;5(6):844.

Wegener HC. Antibiotics in animal feed and their role in resistance development. Curr Opin Microbiol. 2003 Oct;6(5):439-45.

Wesley IV. 1998. Overview: Public Health Significance of *Campylobacter* in Livestock and Poultry, 1998 USAHA (United States Animal Health Association) Proceedings, http://www.usaha.org/speeches/speech98/s98wesle.thml

White P.L., A.R. Baker, and W.O. James. 1997. "Strategies to Control *Salmonella* and *Campylobacter* in Raw Poultry Products", *Rev Sci Tech* Aug;16(2):525-41.

White PA. Use of prior beliefs in the assignment of causal roles: causal powers versus regularity-based accounts. Mem Cognit. 1995 Mar;23(2):243-54.)

Witte W. Selective pressure by antibiotic use in livestock. Int J Antimicrob Agents. 2000 Nov;16 Suppl 1:S19-24.

WHO, 2002. Joint WHO/FAO Expert Consultation on Risk Assessment of *Campylobacter spp*. in broilers chickens and *Vibrio spp*. in seafood, FAO Regional Office for Asia and the Pacific, Bangkok, Thailand 5 -9 August 2002 http://www.fao.org/es/esn/food/risk_mra_campylobacter_2_en.stm

WHO, 2003. Joint FAO/OIE/WHO Expert Workshop on Non-Human Antimicrobial Usage and Antimicrobial Resistance: Scientific Assessment. http://www.who.int/foodsafety/publications/micro/en/amr.pdf, http://www.who.int/foodsafety/micro/meetings/nov2003/en/

WHO/FAO, 2002. Risk assessments of *Salmonella* in eggs and broiler chickens - Interpretative Summary and Full Report.
http://www.who.int/foodsafety/publications/micro/Salmonella/en/

WHO/FAO, 2003. Hazard Characterization for Pathogens in Food and Water - Guidelines.
http://www.fao.org/documents/show_cdr.asp?url_file=/DOCREP/006/Y4666E/y4666e0b.htm

Williams MD, Schorling JB, Barrett LJ, Dudley SM, Orgel I, Koch WC, Shields DS, Thorson SM, Lohr JA, Guerrant RL. Early treatment of *Campylobacter* jejuni enteritis. Antimicrob Agents Chemother. 1989 Feb;33(2):248-50. Erratum in: Antimicrob Agents Chemother 1989 Jul;33(7):1129.

Williamson J. *Bayesian Nets and Causality Philosophical and Computational Foundations.* Oxford University Press. 2005.

Willems RJL, J Top, N van den Braak, A van Belkum, H Endtz, D Mevius, E Stobberingh, A van den Bogaard, and JDA van Embden. 2000. Host Specificity of Vancomycin-Resistant *Enterococcus faecium. Journal of Infectious Diseases*, 182, 816-23. (http://www.journals.uchicago.edu/JID/journal/home.html)

Willems RJ, Homan W, Top J, van Santen-Verheuvel M, Tribe D, Manzioros X, Gaillard C, Vandenbroucke-Grauls CM, Mascini EM, van Kregten E, van Embden JD, Bonten MJ. Variant esp gene as a marker of a distinct genetic lineage of vancomycin-resistant Enterococcus faecium spreading in hospitals. Lancet. 2001 Mar 17;357(9259):853-5.

Winston LG, Bangsberg DR, Chambers HF 3rd, Felt SC, Rosen JI, Charlebois ED, Wong M, Steele L, Gerberding JL, Perdreau-Remington F. Epidemiology of vancomycin-resistant Enterococcus faecium under a selective isolation policy at an urban county hospital. Am J Infect Control. 2002 Nov;30(7):400-6.

Wistrom J, Norrby SR. Fluoroquinolones and bacterial enteritis, when and for whom? J Antimicrob Chemother. 1995 Jul;36(1):23-39.

Wu TL, Su LH, Chia JH, Kao TM, Chiu CH, Kuo AJ, Sun CF. Molecular epidemiology of nalidixic acid-resistant *Campylobacter* isolates from humans and poultry by pulsed-field gel electrophoresis and flagellin gene analysis. Epidemiol Infect. 2002 Aug;129(1):227-31.

Yeh KM, Siu LK, Chang JC, Chang FY. Vancomycin-resistant enterococcus (VRE) carriage and infection in intensive care units. Microb Drug Resist. 2004 Summer;10(2):177-83.

Yoder, H.W. Jr. 1991. "Mycoplasma Gallisepticum Infection". In Calnek, BW, CW Beard, HJ Barnes, WM Reid, and HW Yoder Jr (eds), *Diseases of Poultry*, 9th edn. Iowa State Univeristy Press, Ames. 198-212.

Yokota F, Thompson KM. Value of information analysis in environmental health risk management decisions: past, present, and future.Risk Anal. 2004 Jun;24(3):635-50.

Zhang NL, 1998. Computational Properties of Two Exact Algorithms for Bayesian Networks (1998). http://citeseer.nj.nec.com/zhang98computational.html

Zhang H, Singer B. *Recursive Partitioning in the Health Sciences.* New York:Springer; 1999.)

Zhao C, Ge B, De Villena J, Sudler R, Yeh E, Zhao S, White DG, Wagner D, Meng J. Prevalence of *Campylobacter* spp., Escherichia coli, and *Salmonella* serovars in retail chicken, turkey, pork, and beef from the Greater Washington, D.C., area. Appl Environ Microbiol. 2001 Dec;67(12):5431-6.

Zhao T, Doyle MP, Fedorka-Cray PJ, Zhao P, Ladely S. Occurrence of *Salmonella enterica* serotype *typhimurium* DT104A in retail ground beef. J Food Prot. 2002 Feb;65(2):403-7.

Index

Early Titles in the
INTERNATIONAL SERIES IN
OPERATIONS RESEARCH & MANAGEMENT SCIENCE
Frederick S. Hillier, Series Editor, *Stanford University*

Saigal/ *A MODERN APPROACH TO LINEAR PROGRAMMING*
Nagurney/ *PROJECTED DYNAMICAL SYSTEMS & VARIATIONAL INEQUALITIES WITH APPLICATIONS*
Padberg & Rijal/ *LOCATION, SCHEDULING, DESIGN AND INTEGER PROGRAMMING*
Vanderbei/ *LINEAR PROGRAMMING*
Jaiswal/ *MILITARY OPERATIONS RESEARCH*
Gal & Greenberg/ *ADVANCES IN SENSITIVITY ANALYSIS & PARAMETRIC PROGRAMMING*
Prabhu/ *FOUNDATIONS OF QUEUEING THEORY*
Fang, Rajasekera & Tsao/ *ENTROPY OPTIMIZATION & MATHEMATICAL PROGRAMMING*
Yu/ *OR IN THE AIRLINE INDUSTRY*
Ho & Tang/ *PRODUCT VARIETY MANAGEMENT*
El-Taha & Stidham/ *SAMPLE-PATH ANALYSIS OF QUEUEING SYSTEMS*
Miettinen/ *NONLINEAR MULTIOBJECTIVE OPTIMIZATION*
Chao & Huntington/ *DESIGNING COMPETITIVE ELECTRICITY MARKETS*
Weglarz/ *PROJECT SCHEDULING: RECENT TRENDS & RESULTS*
Sahin & Polatoglu/ *QUALITY, WARRANTY AND PREVENTIVE MAINTENANCE*
Tavares/ *ADVANCES MODELS FOR PROJECT MANAGEMENT*
Tayur, Ganeshan & Magazine/ *QUANTITATIVE MODELS FOR SUPPLY CHAIN MANAGEMENT*
Weyant, J./ *ENERGY AND ENVIRONMENTAL POLICY MODELING*
Shanthikumar, J.G. & Sumita, U./ *APPLIED PROBABILITY AND STOCHASTIC PROCESSES*
Liu, B. & Esogbue, A.O./ *DECISION CRITERIA AND OPTIMAL INVENTORY PROCESSES*
Gal, T., Stewart, T.J., Hanne, T./ *MULTICRITERIA DECISION MAKING: Advances in MCDM Models, Algorithms, Theory, and Applications*
Fox, B.L./ *STRATEGIES FOR QUASI-MONTE CARLO*
Hall, R.W./ *HANDBOOK OF TRANSPORTATION SCIENCE*
Grassman, W.K./ *COMPUTATIONAL PROBABILITY*
Pomerol, J-C. & Barba-Romero, S./ *MULTICRITERION DECISION IN MANAGEMENT*
Axsäter, S./ *INVENTORY CONTROL*
Wolkowicz, H., Saigal, R., & Vandenberghe, L./ *HANDBOOK OF SEMI-DEFINITE PROGRAMMING: Theory, Algorithms, and Applications*
Hobbs, B.F. & Meier, P./ *ENERGY DECISIONS AND THE ENVIRONMENT: A Guide to the Use of Multicriteria Methods*
Dar-El, E./ *HUMAN LEARNING: From Learning Curves to Learning Organizations*
Armstrong, J.S./ *PRINCIPLES OF FORECASTING: A Handbook for Researchers and Practitioners*
Balsamo, S., Personé, V., & Onvural, R./ *ANALYSIS OF QUEUEING NETWORKS WITH BLOCKING*
Bouyssou, D. et al./ *EVALUATION AND DECISION MODELS: A Critical Perspective*
Hanne, T./ *INTELLIGENT STRATEGIES FOR META MULTIPLE CRITERIA DECISION MAKING*
Saaty, T. & Vargas, L./ *MODELS, METHODS, CONCEPTS and APPLICATIONS OF THE ANALYTIC HIERARCHY PROCESS*
Chatterjee, K. & Samuelson, W./ *GAME THEORY AND BUSINESS APPLICATIONS*
Hobbs, B. et al./ *THE NEXT GENERATION OF ELECTRIC POWER UNIT COMMITMENT MODELS*
Vanderbei, R.J./ *LINEAR PROGRAMMING: Foundations and Extensions, 2nd Ed.*
Kimms, A./ *MATHEMATICAL PROGRAMMING AND FINANCIAL OBJECTIVES FOR SCHEDULING PROJECTS*
Baptiste, P., Le Pape, C. & Nuijten, W./ *CONSTRAINT-BASED SCHEDULING*
Feinberg, E. & Shwartz, A./ *HANDBOOK OF MARKOV DECISION PROCESSES: Methods and Applications*
Ramík, J. & Vlach, M./ *GENERALIZED CONCAVITY IN FUZZY OPTIMIZATION AND DECISION ANALYSIS*
Song, J. & Yao, D./ *SUPPLY CHAIN STRUCTURES: Coordination, Information and Optimization*
Kozan, E. & Ohuchi, A./ *OPERATIONS RESEARCH/ MANAGEMENT SCIENCE AT WORK*

Early Titles in the
INTERNATIONAL SERIES IN
OPERATIONS RESEARCH & MANAGEMENT SCIENCE
(Continued)

Bouyssou et al./ *AIDING DECISIONS WITH MULTIPLE CRITERIA: Essays in Honor of Bernard Roy*
Cox, Louis Anthony, Jr./ *RISK ANALYSIS: Foundations, Models and Methods*
Dror, M., L'Ecuyer, P. & Szidarovszky, F./ *MODELING UNCERTAINTY: An Examination of Stochastic Theory, Methods, and Applications*
Dokuchaev, N./ *DYNAMIC PORTFOLIO STRATEGIES: Quantitative Methods and Empirical Rules for Incomplete Information*
Sarker, R., Mohammadian, M. & Yao, X./ *EVOLUTIONARY OPTIMIZATION*
Demeulemeester, R. & Herroelen, W./ *PROJECT SCHEDULING: A Research Handbook*
Gazis, D.C./ *TRAFFIC THEORY*
Zhu/ *QUANTITATIVE MODELS FOR PERFORMANCE EVALUATION AND BENCHMARKING*
Ehrgott & Gandibleux/ *MULTIPLE CRITERIA OPTIMIZATION: State of the Art Annotated Bibliographical Surveys*
Bienstock/ *Potential Function Methods for Approx. Solving Linear Programming Problems*
Matsatsinis & Siskos/ *INTELLIGENT SUPPORT SYSTEMS FOR MARKETING DECISIONS*
Alpern & Gal/ *THE THEORY OF SEARCH GAMES AND RENDEZVOUS*
Hall/ *HANDBOOK OF TRANSPORTATION SCIENCE - 2nd Ed.*
Glover & Kochenberger/ *HANDBOOK OF METAHEURISTICS*
Graves & Ringuest/ *MODELS AND METHODS FOR PROJECT SELECTION: Concepts from Management Science, Finance and Information Technology*
Hassin & Haviv/ *TO QUEUE OR NOT TO QUEUE: Equilibrium Behavior in Queueing Systems*
Gershwin et al/ *ANALYSIS & MODELING OF MANUFACTURING SYSTEMS*

**** A list of the more recent publications in the series is at the front of the book ****