A Risk-Characterization Framework for Decision-Making at the Food and Drug Administration

Committee on Ranking FDA Product Categories Based on Health Consequences, Phase II; National Research Council

ISBN: 0-309-21281-2, 108 pages, 6 x 9, (2011)

PREPUBLICATION COPY

A Risk-Characterization Framework for Decision-Making at the Food and Drug Administration

Committee on Ranking FDA Product Categories Based on Health Consequences, Phase II

Board on Environmental Studies and Toxicology

Division on Earth and Life Studies

National Research Council

Institute of Medicine

NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

THE NATIONAL ACADEMIES PRESS
Washington, D.C.
www.nap.edu

THE NATIONAL ACADEMIES PRESS **500 Fifth Street, NW** **Washington, DC 20001**

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This project was supported by Contract HHSP23320042509XI/HHSP233200800001T between the National Academy of Sciences and the U.S. Department of Health and Human Services. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the view of the organizations or agencies that provided support for this project.

Additional copies of this report are available from

The National Academies Press
500 Fifth Street, NW
Box 285
Washington, DC 20055

800-624-6242
202-334-3313 (in the Washington metropolitan area)
http://www.nap.edu

Printed in the United States of America.

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Charles M. Vest is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. Charles M. Vest are chair and vice chair, respectively, of the National Research Council.

www.national-academies.org

COMMITTEE ON RANKING FDA PRODUCT CATEGORIES BASED ON HEALTH CONSEQUENCES, PHASE II

Members

ROBERT S. LAWRENCE (*Chair*), Johns Hopkins Bloomberg School of Public Health, Baltimore, MD
JEFFREY B. BENDER, University of Minnesota, St. Paul
FRANCISCO DIEZ-GONZALEZ, University of Minnesota, St. Paul
KATHRYN M. EDWARDS, Vanderbilt University, Nashville, TN
SUSAN S. ELLENBERG, University of Pennsylvania, Philadelphia
PAUL S. FISCHBECK, Carnegie Mellon University, Pittsburgh, PA
KAREN E. JENNI, Insight Decisions LLC, Denver, CO
HELEN H. JENSEN, Iowa State University, Ames
L. ROBIN KELLER, University of California, Irvine
JAMES D. MCKEAN, Iowa State University, Ames
DAVID O. MELTZER, University of Chicago, Chicago, IL
SANFORD A. MILLER, University of Maryland, College Park
RICHARD PLATT, Harvard Pilgrim Health Care Institute and Harvard Medical School, Boston, MA
JOHN T. WATSON, University of California, San Diego

Staff

ELLEN MANTUS, Project Director
NORMAN GROSSBLATT, Senior Editor
HEIDI MURRAY-SMITH, Program Officer
MIRSADA KARALIC-LONCAREVIC, Manager, Technical Information Center
RADIAH ROSE, Manager, Editorial Projects
KERI SCHAFFER, Research Associate
PANOLA GOLSON, Program Associate

Sponsor

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES

OTHER REPORTS OF THE BOARD ON ENVIRONMENTAL STUDIES AND TOXICOLOGY

Review of the Environmental Protection Agency's Draft IRIS Assessment of Formaldehyde (2011)
The Use of Title 42 Authority at the U.S. Environmental Protection Agency (2010)
Review of the Environmental Protection Agency's Draft IRIS Assessment of Tetrachloroethylene (2010)
Hidden Costs of Energy: Unpriced Consequences of Energy Production and Use (2009)
Contaminated Water Supplies at Camp Lejeune—Assessing Potential Health Effects (2009)
Review of the Federal Strategy for Nanotechnology-Related Environmental, Health, and Safety Research (2009)
Science and Decisions: Advancing Risk Assessment (2009)
Phthalates and Cumulative Risk Assessment: The Tasks Ahead (2008)
Estimating Mortality Risk Reduction and Economic Benefits from Controlling Ozone Air Pollution (2008)
Respiratory Diseases Research at NIOSH (2008)
Evaluating Research Efficiency in the U.S. Environmental Protection Agency (2008)
Hydrology, Ecology, and Fishes of the Klamath River Basin (2008)
Applications of Toxicogenomic Technologies to Predictive Toxicology and Risk Assessment (2007)
Models in Environmental Regulatory Decision Making (2007)
Toxicity Testing in the Twenty-first Century: A Vision and a Strategy (2007)
Sediment Dredging at Superfund Megasites: Assessing the Effectiveness (2007)
Environmental Impacts of Wind-Energy Projects (2007)
Scientific Review of the Proposed Risk Assessment Bulletin from the Office of Management and Budget (2007)
Assessing the Human Health Risks of Trichloroethylene: Key Scientific Issues (2006)
New Source Review for Stationary Sources of Air Pollution (2006)
Human Biomonitoring for Environmental Chemicals (2006)
Health Risks from Dioxin and Related Compounds: Evaluation of the EPA Reassessment (2006)
Fluoride in Drinking Water: A Scientific Review of EPA's Standards (2006)
State and Federal Standards for Mobile-Source Emissions (2006)
Superfund and Mining Megasites—Lessons from the Coeur d'Alene River Basin (2005)
Health Implications of Perchlorate Ingestion (2005)
Air Quality Management in the United States (2004)
Endangered and Threatened Species of the Platte River (2004)
Atlantic Salmon in Maine (2004)
Endangered and Threatened Fishes in the Klamath River Basin (2004)
Cumulative Environmental Effects of Alaska North Slope Oil and Gas Development (2003)
Estimating the Public Health Benefits of Proposed Air Pollution Regulations (2002)
Biosolids Applied to Land: Advancing Standards and Practices (2002)
The Airliner Cabin Environment and Health of Passengers and Crew (2002)
Arsenic in Drinking Water: 2001 Update (2001)
Evaluating Vehicle Emissions Inspection and Maintenance Programs (2001)
Compensating for Wetland Losses Under the Clean Water Act (2001)
A Risk-Management Strategy for PCB-Contaminated Sediments (2001)
Acute Exposure Guideline Levels for Selected Airborne Chemicals (seven volumes, 2000-2009)
Toxicological Effects of Methylmercury (2000)
Strengthening Science at the U.S. Environmental Protection Agency (2000)
Scientific Frontiers in Developmental Toxicology and Risk Assessment (2000)
Ecological Indicators for the Nation (2000)
Waste Incineration and Public Health (2000)
Hormonally Active Agents in the Environment (1999)
Research Priorities for Airborne Particulate Matter (four volumes, 1998-2004)
The National Research Council's Committee on Toxicology: The First 50 Years (1997)
Carcinogens and Anticarcinogens in the Human Diet (1996)
Upstream: Salmon and Society in the Pacific Northwest (1996)

Science and the Endangered Species Act (1995)
Wetlands: Characteristics and Boundaries (1995)
Biologic Markers (five volumes, 1989-1995)
Science and Judgment in Risk Assessment (1994)
Pesticides in the Diets of Infants and Children (1993)
Dolphins and the Tuna Industry (1992)
Science and the National Parks (1992)
Human Exposure Assessment for Airborne Pollutants (1991)
Rethinking the Ozone Problem in Urban and Regional Air Pollution (1991)
Decline of the Sea Turtles (1990)

Copies of these reports may be ordered from the National Academies Press
(800) 624-6242 or (202) 334-3313
www.nap.edu

Preface

The U.S. Food and Drug Administration (FDA) has regulatory responsibility for a vast number of products—foods, drugs, cosmetics, biologics, veterinary products, medical devices, tobacco, and products that emit radiation—and often faces difficult management decisions as to how to ensure the safety of the products that it regulates. FDA recognized that collecting and evaluating information on the risks posed by the regulated products in a systematic manner would aid in its decision-making process. Consequently, FDA and the Department of Health and Human Services (DHHS) asked the National Research Council (NRC) to develop a conceptual model that could evaluate products or product categories that FDA regulates and provide information on the potential health consequences associated with them. As a result of that request, NRC formed the Committee on Ranking FDA Product Categories Based on Health Consequences.

The project was to be conducted in two phases. For the first phase, the committee was to produce a brief letter report that described a conceptual model that could be used to rank product categories. For the second phase, the committee was to develop, refine, and apply the model to conduct a risk ranking of FDA product categories. The letter report was completed and provided to FDA in February 2009 and is provided in Appendix A. On receipt of the letter report, FDA and DHHS re-evaluated the project and determined that the original scope was too ambitious, that a ranking of products or product categories would be premature, and that the project should be revised to focus more on developing a framework that could be used to evaluate and characterize the public-health consequences associated with FDA-regulated products or product categories in the context of various decision scenarios. Furthermore, FDA and DHHS wanted a framework that would provide a common set of metrics that would enable each center to evaluate the public-health consequences using a common terminology and approach that would allow comparisons within and among disparate programs. The present report reflects the change in scope; it describes the risk-characterization framework proposed by the committee, illustrates the use of that framework with several case studies, and provides the committee's conclusions and recommendations.

This report has been reviewed in draft form by persons chosen for their diverse perspectives and technical expertise in accordance with procedures approved by the NRC's Report Review Committee. The purposes of the independent review are to provide candid and critical comments that will assist the institution in making its published report as sound as possible and to ensure that the report meets institutional standards of objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process. We thank the following for their review of this report: Susan Alpert, Medtronic; Robert L. Buchanan, University of Maryland; Julie A. Caswell, University of Massachusetts; Louis Anthony Cox, Jr., Cox Associates; Robert L. Davis, Kaiser Permanente; James K. Hammitt, Harvard University and Toulouse School of Economics; Lee-Ann Jaykus, North Carolina State University; Gregory M. Paoli, Risk Sciences International, Inc.; Birgit Puschner, University of California, Davis; Jim E. Riviere, North Carolina State University; Joseph V. Rodricks, ENVIRON; and Hugh H. Tilson, University of North Carolina.

Although the reviewers listed above have provided many constructive comments and suggestions, they were not asked to endorse the conclusions or recommendations, nor did they see the final draft of the report before its release. The review of the report was overseen by the review coordinator, Lauren Zeise,

California Environmental Protection Agency, and the review monitor, Chris G. Whipple, ENVIRON. Appointed by NRC, they were responsible for ensuring that an independent examination of the report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of the report rests entirely with the committee and the institution.

The committee gratefully acknowledges staff from several centers at FDA for attending the committee's public session to discuss their current approaches for decision-making at the various centers.

The committee is also grateful for the assistance of the NRC staff in preparing this report. Staff members who contributed to the effort are Ellen Mantus, project director; James Reisa, director of the Board on Environmental Studies and Toxicology; Norman Grossblatt, senior editor; Heidi Murray-Smith, program officer; Mirsada Karalic-Loncarevic, manager of the Technical Information Center; Radiah Rose, manager of editorial projects; Keri Schaffer, research associate; and Panola Golson, program associate.

I would especially like to thank all the members of the committee for their efforts throughout the development of this report.

Robert S. Lawrence, *Chair*
Committee on Ranking FDA Product Categories Based on Health Consequences, Phase II

Contents

TABLES

A Risk-Characterization Framework for Decision-Making at the Food and Drug Administration

Summary

The foundation of the U.S. Food and Drug Administration (FDA) was laid in 1906 with the passage of a law that banned interstate commerce in adulterated or misbranded foods and drugs. Over the last century, FDA has evolved and grown into a government agency that now has over 12,000 employees and regulatory oversight for over $2 trillion in consumer products. FDA has the responsibility to ensure the safety and security of foods, drugs, cosmetics, biologics, veterinary products, medical devices, and products that emit radiation. It must also ensure the efficacy of drugs, biologics, and medical devices and, in June 2009, was given responsibility for regulating tobacco products. The decisions that FDA faces daily can range from determining whether a drug should be approved to allocating resources for inspection of food-production facilities. Decisions often need to be made quickly and on the basis of incomplete information. Given the immensity of its task, FDA recognized that a framework for organizing and evaluating risk-based information in a systematic and consistent manner would be valuable. Accordingly, FDA and the Department of Health and Human Services (DHHS) asked the National Research Council (NRC) to develop a framework that could provide consistent information on health consequences as an aid to decision-making at FDA.

This report, prepared by the Committee on Ranking FDA Product Categories Based on Health Consequences, Phase II, in response to the request from FDA and DHHS, describes a risk-characterization framework that can be used to evaluate and compare the public-health consequences of different decisions concerning a wide variety of products. The framework presented here is intended to complement other risk-based approaches that are in use and under development at FDA, not replace them. The committee recognizes that the public-health-consequence factors highlighted in the framework will seldom, if ever, be the only important considerations in the decision-making process, but they are almost always some of the key considerations.

DECISION-MAKING AT THE FOOD AND DRUG ADMINISTRATION

FDA gave the committee 16 scenarios that highlighted a variety of decisions that FDA regularly faces in which public-health consequences are relevant and for which a systematic and consistent approach for considering risk would be valuable. On the basis of the scenarios, FDA characterized the types of decisions that it faces as mitigation-selection decisions, targeting decisions, and strategic-investment decisions.

- *Mitigation-selection decisions* are those in which FDA must weigh various alternative strategies for addressing a potential health risk. For example, how should FDA balance concerns about the safety of a product with the potential consequences of removing the product from the market?
- *Targeting decisions* are essentially priority-setting or resource-allocation decisions and focus on how particular resources should be allocated among a broad set of products. For example, how should sparse inspection resources be allocated between seafood and fresh produce?

● *Strategic-investment decisions* are longer-term internal decisions about where FDA should invest its resources to enable better risk-informed decision-making. For example, should FDA invest resources to improve collection of data on the food-supply chain or on medical-device surveillance?

The committee notes that there are other ways of categorizing decisions, and some decisions that are within FDA's authority are difficult to fit within the three categories defined here. However, for purposes of developing a decision-focused risk-characterization framework for FDA, the committee adopted FDA's categorization of decisions.

THE RISK-CHARACTERIZATION FRAMEWORK

Health consequences are a subset of the larger array of factors that must be considered for any given problem. Because such factors loom large in most FDA decisions, they constitute a reasonable place to start the process of developing a decision framework. The framework offered here builds on the substantial amount of work that has been done on methods for estimating the human-health consequences associated with various risks, hazards, and decisions. It provides a common language for describing potential public-health consequences of decisions, is designed to have wide applicability among all FDA centers, and draws extensively on the well-vetted risk literature to define the relevant health dimensions for FDA decision-making.

The process is straightforward and involves three steps:

● *Step 1.* Identify and define the decision context: What decision options are being considered? What are the appropriate end points to evaluate and compare?

● *Step 2.* Estimate or characterize the public-health consequences of each option by using the risk attributes that are described below. The values of the risk attributes should be summarized in a table to facilitate comparison of the options.

● *Step 3.* Use the completed characterization as a way to compare decision options and to communicate their public-health consequences within the agency, to decision-makers, and to the public; use the comparison with other decision-relevant information to make informed decisions.

Although the steps can be easily articulated, they involve thought and effort to complete. The framework is not a cookbook and will require FDA to exercise judgment in how it is used. Completing the attribute table (Step 2) may be relatively simple, or it may require substantial research and modeling or even additional data collection and analysis. The decision needs and the available resources should determine how much time and effort should be put into implementing the risk-characterization framework.

THE RISK ATTRIBUTES

Defining a suitable set of risk attributes to characterize the public-health consequences necessary for Step 2 of the framework was a challenging task. Risks are often characterized by a single attribute, such as the number of deaths that could occur as a result of a hazard being evaluated or the probability that an exposed individual will experience an identified adverse effect. However, defining the risk attributes for this framework required recognition of the multidimensional nature of risk.

Consideration of the traditional risk-assessment paradigm gave rise to one set of attributes to characterize health risks. Thus, the committee defined *exposed population*, *mortality*, and *morbidity* as the attributes to use to determine the number, type, and rate of occurrence of adverse health effects that could result from implementation of a particular decision option.

● *Exposed population* is related to the size of the population and the characteristics of the people who are potentially affected by the decision being considered.

● *Mortality* can be described as the number of deaths that will result from the use (or absence) of the product that is the subject of the decision options being evaluated.

● *Morbidity* refers to the illnesses or injuries that are attributable to the decision options being evaluated. This attribute requires a slightly more complex set of metrics that acknowledge differences in severity and duration of a health effect, and the committee therefore proposed the following metrics: *severe adverse health effects* (effects identified as life-threatening, requiring hospitalization, or leading to substantial, persistent, or permanent disability related to impaired organ function), *less severe adverse health effects* (effects that require some level of medical care but are not the more serious effects described above), and *adverse quality-of-life health effects* (effects that may or may not require medical care but have been found to diminish a person's subjective quality of life, such as anxiety, depression, pain, discomfort, and reduced mobility).

Studies of risk perception and public attitudes about risks have consistently shown that although numbers of deaths and illnesses or injuries matter, so do other factors, such as whether a risk is voluntary and how much control a person has over the risks. Thus the committee identified a second set of risk attributes on the basis of the risk-ranking and risk-perception literature that would be applicable to FDA decision-making. Those attributes are *personal controllability*, *ability to detect adverse health effects*, and *ability to mitigate (or reduce) adverse health effects*.

● *Personal controllability* describes the degree to which a person can eliminate or reduce his or her own risks through voluntary action by avoiding exposure to the risk entirely, by reducing the likelihood that exposure will lead to harm, or by minimizing the effects if they do occur.

● *Ability to detect adverse health effects* refers to the ability of informed institutions to detect population-level adverse effects that result from the use (or absence) of the product that is being considered.

● *Ability to mitigate adverse health effects* refers to the ability of institutions to manage, reduce, or otherwise control any expected or unexpected adverse health effects associated with the product that is being evaluated, assuming that such effects exist and are detected.

The attributes proposed here do not preclude the use of additional decision-specific criteria but do capture the major consequences that should be considered in any public-health-related decision. Although the committee recognizes that FDA often must consider other factors—such as economic, social, and political factors—in addition to the public-health consequences in its decision-making (see Figure S-1), it finds that careful and consistent evaluation of the public-health consequences of various options is an essential component of good decision-making.

UNCERTAINTY AND ESTIMATING CONSEQUENCES

Using a single set of risk attributes with specified metrics to evaluate public-health-related decisions entails substantial complexities. For the wide array of decisions that FDA must make, varied amounts of information will be available to estimate the risk attributes for each decision option. For some attributes, large volumes of data may exist for developing or supporting estimates, such as the mortality risk related to some drug. For others, there may be an array of detailed computer models that can help in projecting estimates. For still others, there may be scant data or models available, and direct assessment of the effects by experts will be the only alternative. It is not within the present committee's scope to evaluate, compare, or recommend specific approaches or models for risk quantification. The committee simply urges FDA to bring the best available data and expertise to bear on the evaluation that would be

consistent with the importance of the decision being evaluated and the time and resources available to complete the assessment.

The assessment, quantification, and communication of the uncertainty associated with the estimates are essential. Although categorical measures—such as "likely," "very unlikely," and "possible"—may make the assessment task more palatable, it also makes it considerably less useful. The very ambiguity that provides comfort makes the task of communicating and comparing uncertainties extremely difficult. The committee recommends that the uncertainty in the estimates be described as quantitatively as possible by using summary measures of a probability distribution that describes the estimate of interest. Specifically, the committee suggests that uncertainty be summarized by the 5th, 50th (median), and 95th percentiles of a probability distribution. Although the percentiles are precisely defined, their cognitive interpretation should not be lost. The 5th percentile represents a value below which the actual value is not likely to fall, and similarly, the 95th percentile represents a value above which the actual value is not likely to fall. With training and practice, an analyst can assess an accurate representation of those percentiles from experts using easily described thought experiments, simple tools, and standard protocols. The committee emphasizes that it is always possible to collect more data and do more analyses to try to develop "better" estimates, but there will always be uncertainty, and decisions often must be made on the basis of existing information. Quantifying what is known and what is not known (uncertainty) is an important way to ensure that decisions are as well informed as possible.

CASE STUDIES: LESSONS LEARNED

The committee applied its risk-characterization framework to four hypothetical decision scenarios, each of which was based on scenarios provided by FDA (see Box S-1). In conducting the case-study exercises, the committee reached several conclusions. First, the committee found that it was possible to characterize different decision options by using the risk attributes and that estimates could be made by using existing data and expert judgment. The judgments that were required were not always easy, and committee members were not always comfortable in making them. In the end, however, the committee concluded that the resulting attribute tables would provide useful, relevant, and sufficiently accurate information to be valuable in decision-making.

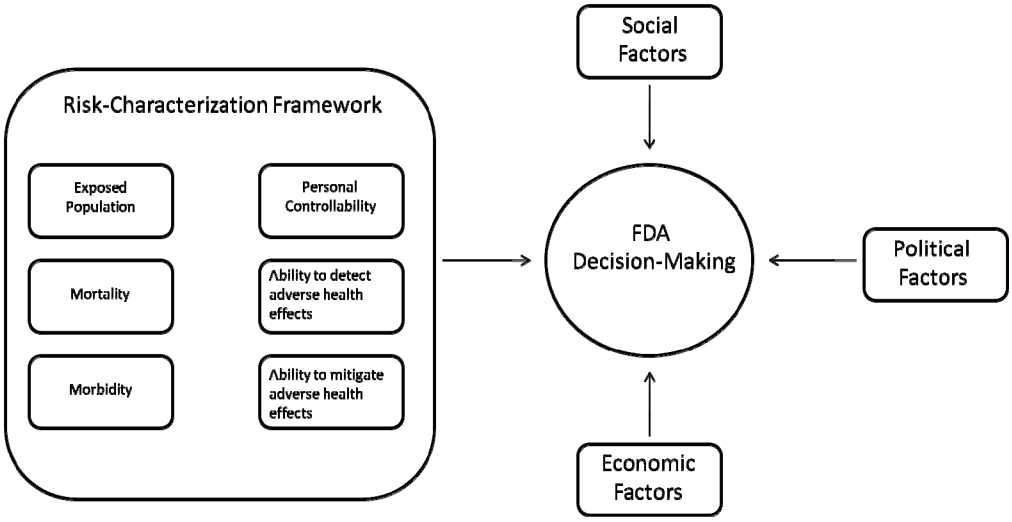

FIGURE S-1 Factors in FDA decision-making.

Second, the value of multiple points of view became evident as the committee developed the case studies. Subject-matter experts were needed to identify and evaluate data relevant to the case studies, and decision analysts were needed to provide the guidance for using the data to estimate the attribute values for the options being compared. The development and analysis of each case study required substantial involvement of both subject matter experts and decision analysts.

Third, the committee found that it was critical in each case to define the decision options to be evaluated and compared clearly so that appropriate risk information for the decision-making process could be obtained. In all cases, analytic reasoning and basic structuring tools, such as influence diagrams, were used to identify the various factors that needed to be considered to develop estimates of the public-health consequences of the alternative decision options.

Fourth, the committee encountered many challenges in finding and interpreting data. In its interactions with FDA, the committee came to recognize that in many cases the agency has a substantial amount of data but the data are not collected, organized, or accessible in a format that is useful to support decisions. The committee emphasizes that simply collecting more data is not necessarily the best use of resources; collecting more relevant data and organizing them so that they are useful in decision-making is the key.

BOX S-1 Description of Case Studies

Mitigation-Selection Decisions. For one case study, the committee considered a hypothetical decision of whether to withdraw a vaccine from the market. It was based on a real-world example in which concerns were raised about a higher-than-expected rate of an adverse effect after vaccination. The committee attempted to look at the situation and the decision options as they were understood at the time when concerns were noted. Although the manufacturer ultimately chose to remove the product from the market, the decision of whether to revoke vaccine licensure is an example of one type of decision that FDA could face in similar circumstances.

Targeting Decisions. For another case study, the committee evaluated the potential public-health consequences of foodborne illness associated with three specific food categories, assuming the current regulatory and inspection regime. The three food categories selected were chosen to highlight products that are inherently different with respect to level of processing, origin, and potential risks. The committee describes how the evaluation could be used directly for ranking or comparing the food categories on the basis of risk or could serve as input into decisions for allocating inspection resources among the food categories to maximize protection of the nation's food supply.

Another case study also examined a *targeting decision*. In this case study, the committee assumed that an analytic laboratory receives many demands from field investigators to test products for a potential contaminant and that extensive testing of any one product with the testing methods available would overwhelm laboratory resources. Before other FDA resources are redirected or outside laboratories are engaged, some sense of the magnitude of the problem must be ascertained. The framework is used to characterize and compare the public-health consequences associated with potential contamination of the given products.

Strategic-Investment Decisions. Because recalls of implanted medical devices are based partly on postmarket-surveillance data, the committee decided to examine a hypothetical situation in which FDA is deciding whether to invest resources in enhanced postmarket-surveillance systems for two specific medical devices—one an established device used extensively across the country and the other an emerging device in limited use. The committee defined the enhanced surveillance systems as furthering the goal of finding types and patterns of unexpected events. Such information should lead to identification of problems in design, implantation processes, clinical interventions, or manufacturing variances; early detection of such problems should lead to improvements. The framework is used to characterize the potential benefits of enhanced surveillance relative to current surveillance approaches.

Fifth, expert judgment and data were inextricably intertwined in how the committee completed each of the case studies. For some case studies, the committee did not have much direct information; in others, adequate data were available. In all cases, assumptions had to be made to interpret the data and complete the attribute table. Although the committee recognizes FDA's strong preference for "data" over "expert judgment" as the basis of any estimates or decisions, it is important to recognize that expert judgment is always present. When decisions must be made immediately, the committee's suggested approach can provide useful information about the public-health consequences of various options in a clear and consistent way on the basis of the best information available at the time the decision must be made. When there is ample time to evaluate and compare decision options, the suggested approach can highlight where additional information would help to differentiate between options (that is, it can help to target information collection) and provide a clear and consistent way to compare the options.

FUTURE DIRECTIONS

To aid decision-making, FDA should consider using the concepts defined by the risk-characterization framework and particularly the risk attributes defined in the present report for discussing risk-related aspects of various decisions. Considering the outcomes of alternative decisions in terms of the attributes identified here will begin to establish consistency in risk vocabulary throughout the agency. As FDA begins to use the risk-characterization framework, it may find that some aspects of the approach need to be modified. Such modifications are entirely appropriate; the approach should evolve to meet FDA's needs as staff gain experience in implementing it.

The committee recognizes that precise predictions of the outcomes of different decisions based on the risk attributes may be difficult to develop. Data may be lacking, and scientists may be uncomfortable in making, or even unwilling to make, the necessary judgments to estimate the risk attributes. However, the committee emphasizes that decisions in which risk information could be valuable are made regularly and recommends that FDA use internal or external experts who are trained in and comfortable with decision analysis, risk assessment, risk management, and specifically the assessment of uncertainties to facilitate the use of the committee's framework.

Changing the organizational culture and its approach to decision-making is a daunting task in any organization. However, FDA is confronted with complex decisions every day, and new approaches are needed to meet the challenges of the future. Using the risk-characterization framework to evaluate the effects of different decisions in terms of the risk attributes described in this report can provide information that is useful in choosing among alternatives.

1

Introduction

The mission of the U.S. Food and Drug Administration (FDA) is to protect public health by ensuring the safety and security of the products that it regulates—foods, drugs, cosmetics, biologics, veterinary products, medical devices, and products that emit radiation. FDA must also ensure the efficacy of drugs, biologics, and medical devices; and in June 2009, it was given responsibility for regulating tobacco products. Given the vast array of products under its regulatory purview, FDA is faced with an enormous task. Globalization of industries regulated by FDA and the complexity of new products and technologies have created new challenges. Expansion of responsibilities and shrinking of resources over the last 20 years have compounded the challenges, and a recent review noted that "FDA is engaged in reactive regulatory priority setting or a fire-fighting regulatory posture" (FDA 2007, p. 4). FDA recognized that its current dilemma could be better addressed if risk information were collected and evaluated in a systematic manner. Consequently, FDA and the Department of Health and Human Services (DHHS) asked the National Research Council (NRC) to convene a committee to develop a framework for evaluating and characterizing the public-health consequences associated with FDA-regulated products or product categories in the context of various decision scenarios. This report, prepared by the Committee on Ranking FDA Product Categories Based on Health Consequences, Phase II, is the response to that request.

THE FOOD AND DRUG ADMINISTRATION AND ITS CENTERS

The modern regulatory functions of FDA date to 1906, when a law that banned interstate commerce in adulterated or misbranded foods and drugs was passed.[1] Since then, over 200 laws have shaped the agency into what exists today. FDA is now a federal agency under the jurisdiction of DHHS and has a staff of over 12,000 and a budget of over $3.3 billion (FDA 2010b). FDA has regulatory oversight of over $2 trillion in consumer products (FDA 2010c) and regulates over 375,000 establishments worldwide (FDA 2007; see Figure 1-1). The five centers described below provide regulatory oversight of foods, drugs, biologics, veterinary products, and medical devices. The information provided is not meant to be a comprehensive review of FDA and all its various activities but simply to highlight the breadth of the agency's responsibilities.

● *Center for Food Safety and Applied Nutrition (CFSAN).* CFSAN has the responsibility for safeguarding the nation's food supply except meat, poultry, and some egg products, which are regulated by the U.S. Department of Agriculture. CFSAN also has regulatory responsibility over dietary supplements and cosmetics and for ensuring that such products are properly labeled. Premarket approvals

[1]For a history of FDA and the laws that have shaped its development, see FDA 2010a.

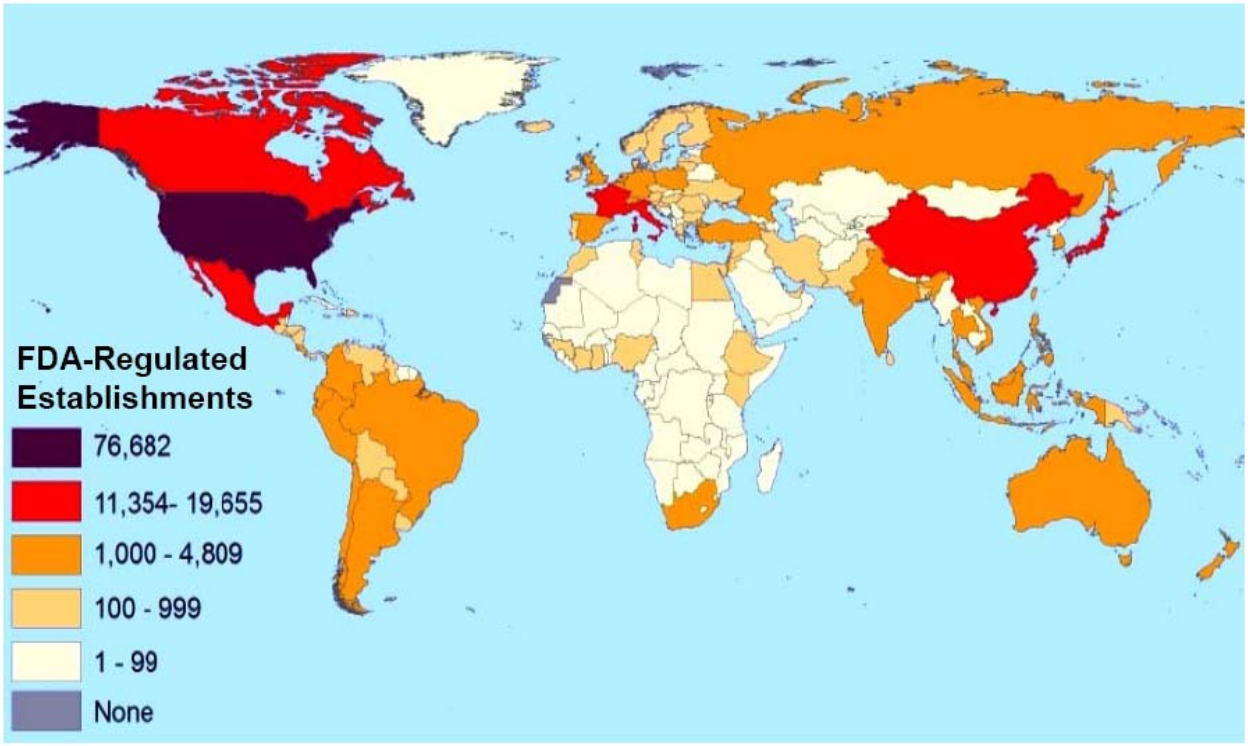

FIGURE 1-1 Worldwide distribution of establishments regulated by the U.S. Food and Drug Administration. Source: FDA 2007, p. 12.

are not required for foods, dietary supplements, or cosmetics, but premarket notifications or approvals are required for food additives and color additives. To ensure product quality, selected manufacturing or processing facilities are inspected in conjunction with FDA's Office of Regulatory Affairs (ORA) to ensure that they are in compliance with good manufacturing practice (GMP). Other activities include monitoring the food supply for contaminants, such as melamine, dioxins, and pesticides; developing new methods to speed detection of contaminated or adulterated food; and maintaining and monitoring a database for adverse-event reporting.

● *Center for Drug Evaluation and Research (CDER).* CDER regulates drug production and distribution and thus has responsibility for ensuring the safety, efficacy, and proper labeling of prescription, generic, and over-the-counter drugs. Drugs require premarket approval, and manufacturing facilities are inspected as part of the approval process. Manufacturing facilities are inspected after drug approval to ensure that GMP is maintained, and drugs are monitored to ensure identity, potency, and content uniformity. To identify unanticipated risks associated with marketed products, CDER collects and evaluates data on drug use and adverse events associated with drug use.

● *Center for Biologics Evaluation and Research (CBER).* CBER ensures the safety, purity, potency, and efficacy of biologics, including vaccines, blood and blood products, cells, tissues, and gene therapies. CBER's processes regarding premarket approval, inspections, and safety surveillance of marketed products are similar to those of CDER, described above. Like CDER, CBER also monitors postmarket problems through an adverse-events reporting database.

● *Center for Veterinary Medicine (CVM).* CVM ensures that veterinary drugs, animal feeds, and pet foods are safe and secure; that veterinary drugs are effective; and that food generated from treated animals does not contain unsafe drug residues. Like human drugs, veterinary drugs require premarket approvals, and ORA inspects manufacturing facilities to ensure compliance with GMP. CVM establishes

standards and monitors animal feed for contaminants. Like the other centers, CVM collects and evaluates data on adverse events associated with veterinary products.

● *Center for Devices and Radiological Health (CDRH).* CDRH ensures that medical devices are safe and effective and that products that emit radiation are safe. CDRH conducts premarket reviews and monitors the manufacturing processes and uses of its regulated products. ORA conducts the premarket and postmarket inspections of manufacturing facilities, performs laboratory analyses, and reviews imports to ensure compliance with FDA standards. Like the other centers, CDRH monitors postmarket problems through an adverse-events reporting database.

THE COMMITTEE'S TASK AND DECISION SCENARIOS FROM THE FOOD AND DRUG ADMINISTRATION

Given its vast responsibilities, its limited resources, and the difficult decisions that it faces daily, FDA recognized that a systematic approach to evaluating the risks associated with its products or product categories would be valuable. Thus, FDA and DHHS asked NRC to convene a committee that could develop a conceptual framework to characterize the public-health consequences associated with its products or product categories in the context of various decision scenarios (see Appendix B for a verbatim statement of task). Such a framework would provide a common set of metrics that would enable each center to evaluate the public-health consequences with a common terminology and a similar approach and would allow comparisons within and among disparate programs.

FDA provided the committee with a set of decision scenarios that illustrate the types of decisions encountered by the various centers except the tobacco program (see Appendix C for the scenarios). The scenarios were provided only as examples of the variety of decisions that FDA regularly faces in which public-health consequences are relevant and in which it might be valuable to consider the consequences systematically and consistently, not as scenarios for the committee to address specifically.

As described by FDA (Bertoni 2010), the scenarios represent three types of decisions faced by the centers:

● *Mitigation-selection decisions* in which FDA must weigh various alternative strategies for addressing a potential health risk. For example, how should FDA balance potential concerns about product safety with the potential consequences of decreased product availability?

● *Targeting decisions* in which FDA must determine which among a broad array of product hazards or potential health benefits should be addressed to obtain the maximum public-health benefit. For example, how should sparse inspection resources be allocated between seafood and fresh produce?

● *Strategic-investment decisions* in which research, data collection, and analytic-tool development can reduce scientific uncertainty and improve FDA's ability to make targeting and mitigation-selection decisions. For example, what public-health concerns should guide a decision about whether to use resources for data collection to improve understanding of food supply-chain safety or for medical-device safety surveillance?

The committee focused on the public-health, scientific, and technologic factors that inform the decisions, although it recognizes that the decisions described above also involve many other factors, including legal and policy considerations.

THE COMMITTEE AND ITS APPROACH TO ITS TASK

Committee members were selected on the basis of expertise in food safety, health economics, medical devices, vaccine safety, pharmacoepidemiology, biostatistics, comparative risk analysis, and decision analysis (see Appendix D for biographic information on the committee). To complete its task,

the committee held six meetings. In a public session during one meeting, FDA staff of various centers discussed with the committee their current approaches and data available for making risk-based decisions. The committee reviewed a variety of literature and made several data requests to the agency so that it could develop case studies to illustrate the use of its framework. The committee did not conduct exhaustive literature searches or reviews to acquire the data for the case studies; it used readily available data or made judgments based on the available data and its multidisciplinary expertise. Thus, the committee emphasizes that all information is illustrative and that the case studies are simply provided as examples of how the committee's framework might be used for the various types of decisions. To be consistent with its statement of task, the committee considered only U.S. users of FDA products and focused on consequences to human health. The committee notes that it was not asked to review or comment on existing decision-making processes at FDA or on the spectrum of risk-based models in development at the agency, and it was not asked to determine how its proposed model would fit within that context.

The committee recognizes that the evaluation of risk is but one factor in the decision-making process. As discussed in the recent report *Enhancing Food Safety: The Role of the Food and Drug Administration*, "risk decision making takes place in a broader social context" (IOM/NRC 2010, p. 75). FDA must also consider such factors as economic constraints and public-health and welfare concerns of all stakeholders. As noted in that report, "it is critical during the information gathering stage to identify which factors will be considered in the decision-making process" (IOM/NRC 2010, p. 75).

ORGANIZATION OF REPORT

The committee's report is organized into seven chapters. Chapter 2 describes the proposed risk-characterization framework and the associated risk attributes. Chapters 3, 4, and 5 provide case studies that illustrate the use of the committee's proposed framework for a mitigation-selection decision, a targeting decision, and a strategic-investment decision, respectively. Chapter 6 provides a case study that examines a targeting decision to set priorities for work that could affect more than one center when choices must be made to allocate agency resources to several pressing needs that arise simultaneously. Chapter 7 describes the lessons learned from developing the case studies and provides the committee's general conclusions and recommendations. The committee's letter report, its statement of task, the decision scenarios, and the committee biographies are provided in appendixes to the report.

REFERENCES

Bertoni, M.J. 2010. Opening Remarks. Presentation at the 5th Meeting on Ranking FDA Product Categories Based on Health Consequences, Phase II, February 3, 2010, Washington, DC.

FDA (U.S. Food and Drug Administration). 2007. FDA Science and Mission at Risk: Report of the Subcommittee on Science and Technology, November 2007. Briefing Information. FDA Science Board Advisory Committee Meeting, December 3, 2007 [online]. Available: http://www.fda.gov/ohrms/dockets/ac/07/briefing/2007-4329b_02_00_index.html [accessed Oct. 25, 2010].

FDA (U.S. Food and Drug Administration). 2010a. About FDA: History. U.S. Food and Drug Administration [online]. Available: http://www.fda.gov/AboutFDA/WhatWeDo/History/default.htm [accessed Oct. 25, 2010]

FDA (U.S. Food and Drug Administration). 2010b. Budget Tables: Summary of Changes FY 2011. FY 2011 Food and Drug Administration Congressional Justification [online]. Available: http://www.fda.gov/downloads/AboutFDA/ReportsManualsForms/Reports/BudgetReports/UCM202309.pdf [accessed Oct. 25, 2010].

FDA (U.S. Food and Drug Administration). 2010c. Executive Summary: Introduction, Mission and Performance Budget Overview FY 2011. FY 2011 Food and Drug Administration Congressional Justification [online].

Available: http://www.fda.gov/downloads/AboutFDA/ReportsManualsForms/Reports/BudgetReports/UCM202307.pdf [accessed Oct. 25, 2010].

IOM/NRC (Institute of Medicine and National Research Council). 2010. Enhancing Food Safety: The Role of the Food and Drug Administration. Washington, DC: National Academies of Sciences.

2

A Risk-Characterization Framework

Difficult decisions are common for the Food and Drug Administration (FDA). Whether it is allocating scarce resources, deciding how to mitigate newly found risks, or deciding what investments in human capital, facilities, data, or analytic methods would be most useful, decision-makers in the FDA often need to integrate data of varied quality, recognize uncertainties, and make trade-offs to arrive at a decision. Public-health and public-safety concerns must be balanced with the economic realities of budgets and the political constraints of imposing new regulations. How a program is presented by the media and understood by the public is an important determinant of the acceptance and success of the program. Science *and* public preferences and perceptions must be considered if one is to understand the potential outcomes of different decision options. To inform the decision-making process, data of differing degrees of quality and robustness must be used and sometimes fed through an array of models of varied sophistication. Expert opinion must be used to interpret the relevance of available data and to solve problems on which available data are weak or nonexistent. Immovable deadlines can thwart uncompromising reliance on the most thorough analysis based on detailed quantitative data for a given decision.

To succeed in such an environment, FDA needs a framework within which alternatives can be defined and evaluated systematically. Although it is beyond the scope of the present study to provide a comprehensive decision-making procedure for FDA, the committee proposes a general framework for thinking about and characterizing the human-health dimensions of FDA decisions. Health consequences are only a subset of the large array of factors that must be considered for any given problem. However, they constitute a reasonable place to start the process of developing a decision framework inasmuch as such factors loom large in most FDA decisions and substantial work has already been done on methods for estimating the human-health consequences associated with various risks, hazards, and decisions.

The framework offered here provides a common language for describing potential public-health consequences of decisions, is designed to have wide applicability among all FDA centers, and draws extensively on the well-vetted risk literature to define the relevant health dimensions for FDA decision-making. This chapter first provides a brief description of the proposed framework and the risk and decision contexts that influenced the committee's approach. Next, the basis and definition of the risk attributes that characterize the framework are provided, and then some approaches for estimating the outcomes of decisions using the risk attributes are described. The chapter concludes with a discussion of how the output of the framework can be used to support decision-making.

THE FRAMEWORK

The risk-characterization framework is designed to be as general as possible while providing consistent risk information in a way that can be used to support the wide variety of decisions that FDA faces. It is intended to supplement and augment other risk-based and risk-informed approaches that are in

use and under development by FDA, not to be a replacement or a one-size-fits-all prescription for conducting all risk-informed decision-making. Indeed, the committee recognizes that the public-health-consequence factors highlighted in this framework will seldom be the *only* important considerations in the decision-making process, but they are almost always some of the key considerations. The U.S. Nuclear Regulatory Commission has also recognized that risks are not the only factors that must be taken into account in regulatory decision-making. It recently embraced the concept of risk-informed decision-making, which it defines as "the use of risk insights, along with other important information, to help in making decisions" (USNRC 2008, page 1-1). The committee's framework focuses on risk information but also recognizes that other information will be relevant for most FDA decisions.

The process is straightforward and involves three steps:

- *Step 1.* Identify and define the decision context: What decision options are being considered? What are the appropriate end points to evaluate and compare?
- *Step 2.* Estimate or characterize the public-health consequences of each option by using the risk attributes that are described below. The values of the risk attributes should be summarized in a table to facilitate comparison of the options.
- *Step 3.* Use the completed characterization as a way to compare decision options and to communicate their public-health consequences within the agency, to decision-makers, and to the public; use the comparison with other decision-relevant information to make informed decisions.

Although the steps can be easily articulated, they involve thought and effort to complete. The framework is not a cookbook and will require FDA to exercise judgment in how it is used. As illustrated in the case studies that are presented in the chapters that follow, successful implementation requires consistency (for example, use of a single set of attributes for every evaluation) and flexibility (for example, specific causes of death or illness that are considered important for a decision can be called out separately in the evaluation and summary). Completing the attribute table (Step 2) may be relatively simple, or it may require substantial research and modeling or even additional data collection and analysis. It should be carried out with whatever level of rigor and effort is necessary and feasible for the decision being considered. Factors to be considered in deciding how much effort to expend on the evaluation of public-health consequences include the timeframe in which the decision must be made and the relative importance of the public-health consequences compared with other key decision-making factors. If a decision must be made quickly and the public-health consequences are less important than other factors, it may be reasonable to complete the attribute table quickly on the basis of available information and judgment alone. However, if the decision options under consideration could lead to substantially different public-health consequences or if public-health consequences are highly uncertain or poorly understood and other factors are considered less important, it may be worth substantial time and effort to develop more precise estimates for the attribute tables. The decision needs and the available resources should determine how much time and effort should be put into implementing the risk-characterization framework.

CONTEXT PROVIDED BY RISK LITERATURE

Science and Decisions: Advancing Risk Assessment (NRC 2009) lays out a "framework for risk-based decision-making" as part of its recommendations for improving risk assessment at the U.S. Environmental Protection Agency (EPA). Although that study was focused specifically on formal risk assessment, it summarized its approach into an overall decision framework that closely matches what the present committee concludes is appropriate for improving risk-based decision-making at FDA. Figure 2-1 illustrates the *Science and Decisions* framework.

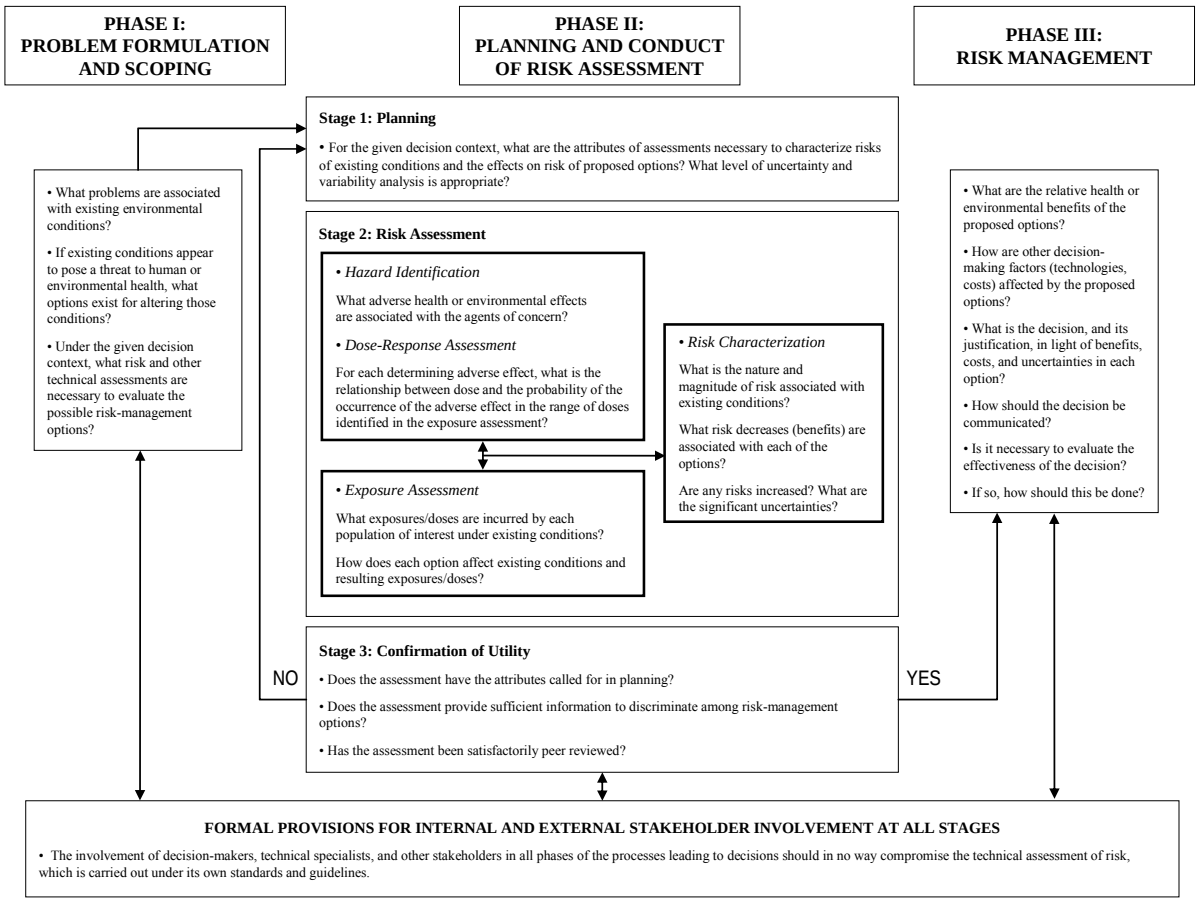

FIGURE 2-1 A framework for risk-based decision-making that maximizes the utility of risk assessment. Source: NRC 2009, p. 243.

The three phases shown in Figure 2-1 correspond to the three steps in the framework listed above, but the emphasis of the present committee is somewhat different from that of the previous study. Specifically, the bulk of *Science and Decisions* focuses on the risk-assessment portion in phase II of the framework, as is highlighted in the figure. The framework proposed here focuses on the risk-characterization portion of the decision structure and adopts a more robust view of the importance of risk characterization in the overall process of making risk-informed decisions.

Early descriptions of the risk-assessment process identify risk characterization simply as a process of combining the exposure and dose-response elements of a risk assessment to summarize and communicate results. In Figure 2-1, for example, risk characterization is placed in its traditional location as the last step in a risk assessment. *Understanding Risk* (NRC 1996, p. 16) laid out a challenge to that narrow interpretation and greatly expanded the scope and role that risk characterization should play in the overall process of making risk-informed decisions:

> We have concluded that the view of risk characterization as a summary is seriously deficient, and we propose a more robust construction. Risk characterization must be seen as an integral part of the entire process of risk decision making: what is needed for successful characterization of risk must be considered at the very beginning of the process and must to a great extent drive risk analysis. If risk characterization is to fulfill its purpose, it must (1) be decision-driven, (2)

> recognize all significant concerns, (3) reflect both analysis and deliberation, with appropriate input from the interested and affected parties, and (4) be appropriate to the decision.

The present committee adopts the broader view of risk characterization and its relationship to risk-informed decision-making.

The committee's approach was also influenced by comparative risk analysis (CRA), which was first defined by EPA in *Unfinished Business* (EPA 1987). That report embraces the multidimensional nature of risk and ranks 31 identified environmental threats according to four attributes: cancer risk, noncancer risk, ecologic risk, and "welfare" effects. In a follow-up report, *Reducing Risk* (EPA SAB 1990), EPA's Science Advisory Board endorsed the broad CRA approach and as a result spawned many applications of CRA at the office, region, state, and local levels (Minard 1996; Jones 1997). Those early CRA efforts led to questions about how best to facilitate comparisons and identify useful attributes for characterizing risks or risk-reduction opportunities.

Progress on CRA method development continues, although its use remains relatively limited. In February 1994, a workshop organized by Resources for the Future for the president's Office of Science and Technology Policy brought together researchers in CRA with the goal of developing a systematic process for comparing risks among different federal agencies (Davies 1996). As part of that work, researchers at Carnegie Mellon University developed a framework for ranking risks that included both quantitative and qualitative measures of relevant programmatic attributes (Fischhoff 1995; Morgan et al. 1996). They included health-effect measures (such as morbidity and mortality) and psychometric measures that research shows play an important role in the evaluation of risks (such as fairness, scientific understanding, and uncertainty). That work spawned a series of research projects and papers that refined and applied the framework (see, for example, Morgan et al. 1999, 2001; Long and Fischhoff 2000; Morgan et al. 2000; DeKay et al. 2001; Florig et al. 2001; Willis et al. 2004, 2005; Fischhoff 2006; Gutiérrez et al. 2006; Bronfman et al. 2007, 2008a,b), including a discussion directly related to food safety (DeKay et al. 2005).

The framework defined here builds on that work while embracing the notion that risk characterization should be decision-focused rather than restricted to ranking risks. The framework focuses on describing the potential effects of alternative *decisions* on health rather than on comparing different health and environmental hazards. It also focuses on the identification and use of a clear and consistent set of risk attributes relevant to public health to describe the effects of alternative decisions.

Although the health outcomes of alternative decisions are not the only factors influencing regulatory decision-making, such information is often highly useful in weighing the merits of alternative decisions. For example, in deciding whether to withdraw approval for a product for which a new adverse health effect has been identified, knowledge of the extent, likelihood, and severity of the newly identified adverse health effect is important, but so is knowledge of the product's benefits and of the effects that are likely to occur if the product is no longer available. An important question for the present committee was how to define a set of attributes that would be robust enough such that risk information relevant to the broad array of decisions that FDA faces could be adequately captured and used consistently for risk-informed decision-making.

DECISION CONTEXT

As noted in Chapter 1, FDA provided the committee with 16 decision scenarios (see Appendix C) and used them to characterize the types of decisions that it faces. FDA asked the committee to consider three types of decisions: mitigation-selection decisions, targeting decisions, and strategic-investment decisions (Bertoni 2010). For purposes of developing a decision-focused risk-characterization framework for FDA, the committee adopted FDA's categorization of decisions.

Other ways of categorizing decisions are possible, and some decisions that are within FDA's authority are difficult to fit within the three categories defined here. For example, decisions about setting

or modifying regulatory standards or establishing certification standards have the potential to affect numerous FDA-regulated products simultaneously, and the effects of such standards depend not only on the standards but on the response of the regulated industry to the standards and on the effectiveness of enforcement actions. The committee has not explicitly addressed the standard-setting decisions in the present report, although the concepts presented here could be extended to address such decisions. Others have addressed the use of a comprehensive risk perspective in setting standards (Fischhoff 1984), including consideration of when standard-setting is preferable to case-by-case decision-making.

Mitigation-Selection Decisions

Mitigation-selection decisions require FDA to choose among two or more options that are available to reduce or mitigate identified risks. The first step in applying the framework to such decisions is to identify and specify the mitigation decision and the decision options to be evaluated and compared. For example, in the case study described in Chapter 3, the decision context is a hypothetical situation in which a vaccine side effect is believed to be occurring at a rate higher than expected. Two decision options are considered—remove or do not remove the vaccine from the market. In this case study, the committee determined that it was important to consider the health consequences of both the vaccine and the underlying disease that the vaccine is intended to prevent.

The decision scenarios provided to the committee by FDA included several additional examples of mitigation-selection decisions, some of which may require evaluating and comparing a larger number or greater variety of options beyond simple binary choices of recall or not. For example, for product-recall decisions, there might be different levels or types of recall that could be executed, each of which could lead to different health consequences; there might be mitigation options that combine recall of unused product with other risk-reduction options for product already in use. For risk-mitigation decisions with several options, each of the viable risk-mitigation approaches needs to be specified, and the consequences of each must be estimated by using the risk attributes defined below. More generally, almost any decision that results in a change in product availability or the extent of use of a particular product could be characterized as a mitigation-selection decision; the outcomes will differ in the degree of use of the product, and the resulting health consequences can be characterized and compared by using the framework proposed here.

Targeting Decisions

Targeting decisions as described by FDA are essentially priority-setting or resource-allocation decisions. They appear to be made primarily within a program and focus on how a particular resource, such as inspection capability, should be allocated among a broad set of products or product categories. In this type of decision, the options or alternatives theoretically available to FDA are vast: virtually any amount of a resource could be allocated to any of a subset of the identified products or product categories, and the only constraint would be total resource availability. The large array of options will need to be narrowed judiciously before the effects of the resulting decisions can be evaluated in detail.

For each product, product category, or other item for which resources might be targeted, FDA will need to identify the resources that it is considering allocating. In some cases, the decision may be an "all or none" decision, such as a decision to inspect a facility or not to inspect it. In others, it may be a particular level of effort and resources, such as a specific rigor of inspection of a facility. Generally, each potential target of a resource allocation would require evaluation of at least two decisions: allocating "x" level of resources vs allocating "y" level of resources. The difference between the values of the risk attributes of an allocation of "x" vs "y" defines the public-health benefit of allocation of resources at those levels to each product or product category.

As noted by FDA (Bertoni 2010), targeting decisions can be seen as similar to the risk-ranking questions that are historically the main focus of CRA studies and that remain of interest in some FDA centers. For example, the Center for Food Safety and Applied Nutrition (CFSAN) has developed a "Fresh Produce Risk Ranking Tool," which can be used to rank produce-pathogen pairs in order of risk (FoodRisk.org 2011). Similarly, a detailed risk assessment of *Listeria monocytogenes* conducted in 2003 led to a ranking of ready-to-eat foods by the likelihood and frequency of illnesses (FDA/USDA 2003).

It is unclear, however, how risk ranking can or should be used to make risk-informed decisions. When different allocations of resources are expected to be equally effective in reducing the risks associated with the product or product category in question, a ranking of potential targets according to unreduced risk may provide sufficient decision support. For example, in making decisions about what facilities to inspect in a specific year, CFSAN may be able to use a risk ranking of facilities in their current conditions and operations with the assumption that additional resources will be equally effective in reducing the risks at any facility. Given that assumption, resources should be focused on the highest-risk facilities. To create such a risk ranking, FDA would need to estimate the public-health consequences of the food products from each facility that is a candidate for inspection—or, more generally, for each potential target of the resource allocation—on the basis of its best assessment of the current conditions and operations of the facilities. The resulting ranked list could then be used as the basis of resource-allocation decisions.

However, the assumption of equal effectiveness is not often appropriate in real decision problems. If the situation is more complicated—that is, if the risks associated with some products or product categories being evaluated can be reduced much more effectively or at much lower cost than the risks associated with others—a narrowly focused risk ranking will not be an appropriate tool for making the decisions. In such situations, FDA should define the decision options more explicitly, as discussed above. For example, the risk associated with a food from a particular facility could be determined for two inspection regimes. The basis of the status quo inspection procedures could be extracted from an existing risk-ranking exercise. Then FDA would also need to evaluate the public-health consequences associated with the food from the facility if a decision were made to allocate a different level of inspection resources to the facility; this evaluation would include changes in any risks affected by the different inspection. The same exercise would then need to be performed for any other food-facility combination that is being considered in the resource-allocation decision. That approach would lead to a "risk-change" ranking rather than a risk ranking and would enable FDA to take into account the possibility that some inspections are more effective than others in reducing risks. If resources are to be reallocated from existing inspection programs to new programs, the increases in risk due to the reduction in funding in one must be balanced against the decreases in risk due to increased funding in the other.

In Chapter 4, the committee illustrates the use of the framework to evaluate three food categories for their current health effects. The results of the evaluation could be used to rank the food categories by risk, and the ranking could be used, subject to the limitations described above, to support a targeting decision for a hypothetical example of food-safety inspections. In Chapter 6, the committee illustrates a targeting decision that involves a more explicit comparison of the effects of different allocations of resources.

Strategic-Investment Decisions

Strategic-investment decisions are longer-term internal decisions about where FDA should invest its resources to enable better risk-informed decision-making. For example, additional research can be conducted, more data collected and analyzed, and new tools developed to increase understanding of and reduce uncertainties about a wide array of potential risks. Investments in such activities typically enable better evaluations and more informed decisions, so they provide value to FDA and the American public. However, they are tied only indirectly to the more common metrics of public health, and it can be difficult to demonstrate their benefits and to decide which strategic investments are most worth while.

The decision-analytic concept of *value of information* can be used to evaluate the potential benefits of strategic-investment decisions about data or information collection. This concept is that new information has decision-relevant value only if it could lead to actions different from the actions that would be taken without the information. That is, information and strategic investments to collect it have value only if they have the potential to change decisions and thus potentially improve outcomes. For example, if a physician orders a diagnostic test but would recommend the same treatment regardless of the test results, the information has no decision-relevant value. In contrast, if the treatment recommendation depends on the diagnosis and the test results will be used to differentiate among diagnoses, the information potentially has decision-relevant value. The risk-characterization framework can provide some of the key elements necessary for a formal value-of-information analysis: definition of a decision context (Step 1) and characterization of decision outcomes (Step 2). Some extensions to the framework would be necessary to implement a value-of-information analysis fully, and those are explored in several of the case studies that follow.

In the case study in Chapter 5, the committee uses its approach to evaluate the public-health consequences of a strategic-investment decision to enhance postmarket surveillance of two medical devices. Application of the committee's proposed approach to this decision category is the most challenging both because identifying the strategic-investment decision options to be evaluated is complicated and because evaluating the effects of long-term strategic investments on public health is difficult. The case study allows some comparisons to be made but probably represents a narrower scope of strategic-investment options than what FDA would consider.

CHARACTERIZING THE HEALTH CONSEQUENCES OF DECISION OPTIONS

In the risk-assessment paradigm, risks have traditionally been characterized by a single attribute, such as the number of deaths that could occur as a result of the hazard being evaluated or the probability that an exposed individual will experience an identified adverse effect. Several complications arise with that simple characterization of risks. Multiple outcomes, such as illnesses and deaths, are often of interest. Uncertainty in the outcomes makes simple characterization and reporting problematic: reporting only expected fatalities will obscure information about the other possible outcomes that are important to decision-makers and policy-makers. Providing only single values for health outcomes may convey far greater certainty than is appropriate. Characterizing the public-health consequences requires recognizing and accommodating the multidimensional nature of risk and the uncertainties involved in estimating the outcomes.

The Many Dimensions of Risk

The comparison of alternative decisions may seem easier when a single metric is used and when the harm and the benefit are measured with the same metric. For example, in deciding whether the association of clozapine with the potential fatal side effect of agranulocytosis is sufficient to warrant keeping clozapine off the market, the number of deaths caused by clozapine-induced agranulocytosis can be directly compared with the number of suicides that clozapine would be expected to prevent. However, those estimates may differ greatly in their uncertainty, and other benefits of clozapine, such as improved ability to function, would also have to be considered in such a decision.

When the harm and the benefit of alternative decisions cannot be measured with the same metric, decision-making is less straightforward, and the importance of informing decision-makers of the various harms and benefits becomes more important. For example, infliximab is highly effective in reducing pain and improving function in people who have rheumatoid arthritis, but its immunosuppressive properties may permit serious and possibly fatal infections to emerge. Deciding whether the risk of a serious adverse effect outweighs the benefits in increased mobility and quality of life is not a technical or

scientific question but rather a question of personal and societal values. An integral part of the committee's proposed approach is that it characterizes various effects explicitly so that decision-makers can make informed decisions that account for them rather than combining them into a single metric based on implicit weightings (see Box 2-1).

The risk-assessment paradigm gives rise to one set of attributes for characterizing the health risks associated with FDA-regulated products or FDA decisions: factors that are used to determine the number, type, and rate of occurrence of adverse health effects (including deaths) that could result from implementation of a particular decision option. Those attributes are *exposed population*, *mortality*, and *morbidity*, each described in more detail below.

Studies of risk perception and public attitudes about risks have consistently shown that although the mortality and morbidity components in risk estimation are important, they are not the only things that people care about when they think about risks and about risk acceptability (Slovic 1992). Numbers of deaths and illnesses or injuries matter, but so do other factors, such as whether a risk is voluntary, how much control a person has over risks, and whether the hazard being considered has the potential to lead to a large number of simultaneous deaths. The list of factors hypothesized to be important in understanding risk is long (see Appendix E). Relatively few risk attributes have been studied in detail, but the available studies show that the various factors are highly correlated on a relatively small number of dimensions, and it has become common to refer to these dimensions as reflecting key "factors" that characterize risk perceptions (Slovic 1992; NRC 1996). The first dimension captures the quantitative aspects, such as the number and type of adverse health outcomes. The second dimension is roughly characterized by the degree of knowledge about the hazards or risks and how well they are understood. The third dimension is characterized by variables that are less easily summarized by a single category: whether a risk is voluntary, how much control an exposed person has over it, the ease with which it can be reduced, and whether it is catastrophic (that is, can lead to multiple simultaneous deaths or injuries).

Understanding Risk urged that risk characterization explicitly include consideration of those additional factors and described a number of ways in which they could be taken into account in decision processes (NRC 1996, p. 65). That approach has also been embraced as a key part of CRA and in the risk-ranking studies mentioned above. The risk-ranking studies include both quantitative information about the number of deaths and injuries and more qualitative information reflecting some of the variables identified in the risk-perception literature. Little general guidance, however, has been offered about what risk attributes might be widely appropriate. Investigators in the studies cited above typically acknowledge that attribute selection is complicated and then choose attributes that they have concluded are appropriate for their particular study context. Jenni (1997) used three methods for assessing the importance of risk attributes for two types of risk-policy decisions and found no universal attributes. The first method used a traditional psychometric risk-perception approach in which subjects evaluated an array of hazards in terms of perceived risk and acceptability of those risks and rated the hazards by using various attributes. The correlations between risk acceptability and attribute ratings were used to determine attribute importance. The results were compared with those of a study in which subjects provided direct ratings of the importance of considering each attribute in risk-informed decision-making and a third study in which the relative importance of various attributes was derived from hazard comparisons. For the hazard comparisons, subjects reviewed pairs of hazards described in terms of their consequences on multiple attributes and stated which hazards were more important to address; this led to an indirect assessment of attribute importance. Each study used two distinct hazard domains (general technologic risks and risks to students in school). The different hazard domains and the different assessment methods led to different results on attribute importance; that finding makes it difficult to argue that one definitive set of risk attributes is relevant for all decision-makers and all decision types.

Ideally, attributes appropriate for characterizing risks in a specific context would be identified jointly by the decision-making organization and affected stakeholder groups through an iterative analytic-deliberative process (NRC 1996). In the present study, however, FDA is interested in a method that provides a common language for characterizing the risks of public-health consequences across a broad

BOX 2-1 Summary Measures vs Detailed Characterization of Public-Health Consequences

In the public health literature, it has become common to measure and communicate the burden of disease by using a summary measure that combines mortality and morbidity into a single value. Some measures, such as disability adjusted life years (DALYs), have been used to evaluate the global burden of disease (Lopez et al. 2006). Others, such as quality adjusted life years (QALYs), have been promoted as tools for comparing the cost effectiveness of various medical interventions (Broome 1993) or risk reducing regulatory strategies (IOM 2006). Each summary measure necessarily embeds a set of value judgments about different levels of health impairment, and each provides a narrow, rather than a robust, characterization of the risk. The committee finds that a richer characterization of the public-health consequences of decision options is needed for routine comparison of decision options. As described in *Improving Risk Communication*, "reducing different kinds of hazard to a common metric (such as number of fatalities per year) and presenting comparisons only on that metric have great potential to produce misunderstanding and conflict and to engender mistrust of expertise" (NRC 1989, p. 52).

array of potential decision types and product categories, not one that is necessarily decision-specific. To address that need, the committee identified a second set of risk attributes on the basis of the risk-ranking and risk-perception literature: *personal controllability, ability to detect adverse health effects, and the ability to mitigate (or reduce) adverse effects.* Because the framework focuses on supporting risk-informed decisions, those attributes, which are described in detail below, are driven more strongly by studies of the acceptability of risk (see, for example, Slovic 1987) and the direct assessment of attribute importance (Jenni 1997) than by the factors found to correlate with judgment of perceived risk that have been more widely described (see, for example, Slovic 1992).

Risk Attributes for Characterizing the Public-Health Consequences of Decisions

Exposed Population

Exposed population is defined here by using two metrics related to the population size and the characteristics of the people who are potentially affected by the decision being considered. The first metric is the number of people in the United States whose health could be affected in a specified timeframe (annually in this report) by the decision being considered. Determining the exposed population requires judgment and consideration of the decision context. For example, the exposed population for decisions that affect the availability or use of a particular product may consist of all people who use (or would use if it were available) the product of interest. The second metric is intended to capture groups or populations of special concern (if any) that may have higher exposure or be more sensitive to potential adverse effects of that exposure than the general population. For example, if the exposed population consisted disproportionately of children, that fact should be included in the summary. Any populations of special concern to FDA should be highlighted. Recommendations from the Institute of Medicine (IOM 2009) on priorities for comparative-effectiveness research highlight studies that focus on different populations and suggest that the following groups may be of particular interest: children, the elderly, ethnic populations, people who have disabilities or chronic diseases, pregnant women, and people who are immunosuppressed. Products that are used disproportionately by members of one or more of those groups should be noted.

Mortality

Mortality is defined for this study as the number of deaths that will result annually from the use (or the absence) of products that are the subject of the decision options being evaluated. The number of deaths can be combined with the number of people exposed to yield a mortality rate if such a metric is of interest. Keeping the two attributes (exposed population and number of deaths) separate, however, provides more information than the mortality rate alone and will allow FDA and other audiences to distinguish between risks that are broadly distributed among the population from risks that are concentrated in a smaller group. It will also allow readers to determine whether a particular group is disproportionately at risk or disproportionately affected by the potential decision options.

Morbidity

Morbidity refers to illnesses or injuries and requires a slightly more complex set of metrics that acknowledge differences in the severity and duration of a health effect. Three metrics are suggested as a way to summarize health effects of different severity and duration that may be viewed or valued differently. The categories are intended to be mutually exclusive.

- *Severe adverse heath effects*—health effects identified as life-threatening, requiring hospitalization, or leading to substantial, persistent, or permanent disability related to impaired organ function.
- *Less severe adverse health effects*—health effects that require some level of medical care but are not the more serious effects described above. The annual number of office visits is one measure of the prevalence of less severe adverse health effects. Pathophysiologic changes in this category are generally reversible.
- *Adverse quality-of-life health effects*—a suite of other effects that may or may not require medical care but have been found to diminish a person's subjective quality of life. This category includes any of the effects described by the EuroQoL system: anxiety or depression, pain or discomfort, inability to participate in usual activities, inability to care for oneself, and reduced mobility (EuroQoL 1990; AHRQ 2005).

The risks associated with the decision options being evaluated should be characterized by the number of people who would suffer from each of the types of adverse health effects annually.

Personal Controllability

The attribute *personal controllability* describes the degree to which a person can eliminate or reduce his or her own risks through voluntary action by avoiding exposure to the risk entirely, by reducing the likelihood that exposure will lead to harm, or by minimizing the effects if they do occur. Personal controllability is related to several risk variables that have been found to be important in the risk-perception, risk-acceptability, and attribute-importance studies discussed above, particularly voluntariness and controllability.

Three critical factors determine the ability of a person to control or limit his or her risks. First, people must be aware that risks exist and that they are potentially exposed to harm. Second, options must be available for avoiding, eliminating, or reducing the risk of harm associated with the product being considered. Third, the potentially exposed person must have knowledge of the options and the ability to choose one. Note that this attribute does not necessarily imply that the exposed person will take steps to minimize the risks, only that he or she would be able to do so.

Measuring or estimating the personal controllability of risks is not trivial: knowledge of risks and the ability to choose from among options can vary substantially across the population of exposed people. Some people will have knowledge of the risks but will not have (or will think that they do not have) the ability to choose from among options, and some people will not have knowledge of the risks or options but would have the ability to choose from among options if they knew about them. Neither group, however, has any degree of personal controllability. Therefore, the appropriate metric for personal controllability might be the percentage of the exposed population that has both the appropriate knowledge and the ability to exercise risk-reducing strategies to be able to control their own risks. The case studies in the chapters that follow illustrate how this metric can be estimated.

Ability to Detect Adverse Health Effects

Ability to detect adverse health effects refers to the ability of informed institutions to detect population-level adverse effects that result from the use (or absence) of the product that is being considered. Here, *institution* refers not only to FDA but to any centralized group that has a role in oversight, distribution, or application of the product being evaluated. Thus, it includes other public-health agencies and organizations, hospitals, pharmaceutical companies, food distributors, cosmetics producers, and others. The focus is on the ability to detect adverse effects that are occurring in the population at rates greater than expected because of use (or absence) of the product being evaluated, that is, to notice adverse health effects and determine that their cause is related to the product being evaluated.

This attribute is related to a different set of risk variables from those associated with personal controllability. Specifically, it is proposed as a proxy for a variety of factors related to how well a risk or hazard is understood (for example, the quality of scientific understanding of the risk or hazard) and the difficulties surrounding detection of the adverse effects associated with the risk or hazard (for example, a long latency between exposure and effect or other factors would make it difficult to associate the adverse effect with the risk or hazard). Although the importance of some of those factors may be captured by adequately representing the uncertainty in the effects (discussed in the next section), this attribute captures the less measurable concerns about whether and how quickly risks will be recognized.

There are several ways to consider and characterize the ability to detect population-level adverse effects, and different metrics may be necessary for different decision contexts. For example, in the case studies that follow, the ability to detect botulism poisoning from canned foods is considered. Because botulism poisoning is extremely serious and rare, and there is a requirement to report every incident, even a single case of botulism poisoning from canned foods is likely to be detected. For those types of decision problems, detectability might be characterized by the probability that a problem will be recognized or by the percentage of adverse effects that will be recognized. In another case study, the ability to detect a side effect of a specific vaccine is considered. The side effect is a medical condition that occurs with moderate frequency in the population in the absence of the vaccine being evaluated. Although a slight increase in the incidence of that condition might lead to a relatively high number of illnesses, it may be difficult to distinguish the change from the background rate, so individual occurrences of the side effect are easy to recognize, but the systemic increase in the rate of occurrence due to the vaccine is more difficult to recognize. In such a case, it may be more useful to estimate the detectability of the problem by estimating how prevalent unexpected adverse health effects would have to be (for example, would have to affect 1 in 10,000 of those exposed) or how much more frequently they would have to occur (for example, would need an effect rate 20% higher than the baseline) for them to be detected and correctly attributed.

Ability to Mitigate Adverse Health Effects

Ability to mitigate adverse health effects refers to the ability of institutions to manage, reduce, or otherwise control any expected or unexpected adverse health effects associated with the product that is

being evaluated, assuming that such effects exist and are detected. When institutions detect that a problem is occurring at a rate higher than expected or leading to more severe effects than expected, there may be actions that those institutions can take to reduce the severity of the problem. For example, a product-recall decision, if successful, reduces further exposure to the product and thereby reduces the number of people who will suffer adverse effects. Another mitigation measure could be recommending that patients who have medical implants take antibiotics before dental work to reduce the likelihood of infection and thereby mitigate a potentially serious health effect. The attribute is related to controllability, reversibility, and the ease with which risks can be reduced—all of which have been shown to be important in the risk-perception, risk-acceptability, and attribute-importance studies discussed above. The ability of a person to mitigate his or her own risk is captured in the personal-controllability attribute, but the attribute being discussed here focuses on whether large-scale institutional actions are available to reduce the extent or severity of adverse effects.

The committee proposes that the ability to mitigate adverse effects be characterized by the probability that an informed institution will be able to reduce (mitigate) adverse health effects associated with the product being evaluated if such a problem is known to exist or, alternatively, by the estimated percentage of potentially affected people whose risks can be reduced (mitigated) by institutional actions. For example, the effectiveness of a recall can be used to estimate the percentage of people whose risks can be reduced; thus, if the institution can remove 90% of its product from the market, it can prevent exposure of and reduce the risk to 90% of product users. Another example would be the effectiveness of a treatment for the problem that has resulted from exposure to the product; for example, 90% of the people suffering from the ill effects can be successfully treated.

Summary of Attributes and Metrics

Table 2-1 summarizes the six risk attributes and the metrics to be used to evaluate the public-health consequences of alternative decisions.

The risk attributes in Table 2-1 focus on providing information relevant to the public-health consequences of the various types of decisions described by FDA. The committee recognizes that FDA must consider other factors—such as economic, social, and political factors and decision-specific factors—in addition to public-health consequences in its decision-making and that some factors commonly discussed are not explicitly included in the list above (see Box 2-2). The attributes proposed here are not intended to preclude the use of additional decision-specific criteria but to capture the major consequences that should be considered in any public-health-related decision. The committee considers careful and consistent evaluation of the public-health consequences of various options to be an essential component of good decision-making.

Using the Risk Attributes to Characterize Health Consequences

Using a single set of risk attributes with specified metrics to characterize the public-health consequences of different decisions entails substantial complexities. First is the wide array of decisions that FDA must make, as described in Chapter 1 and elaborated in the discussion of decision context above. Second is that the availability of data to support such estimates varies greatly across the spectrum of public-health-related decisions that FDA makes. Third is the need to characterize the uncertainty in the consequences. Fourth is the level of comfort with making estimates on the basis of incomplete information and substantial uncertainties, which varies greatly within the agency and among individual scientists. Fifth is accurately communicating the attribute values and their uncertainties to decision-makers and policy-makers without introducing biases, such as overconfidence and false precision.

TABLE 2-1 Summary of Attributes and Metrics

Attribute	Metrics
Exposed population	a) Number of people in the United States directly affected each year by the product or resource that is the subject of the decision being considered
	b) List of populations of concern that use the product or resource in a year
Mortality	Number of deaths in a year attributable to the product or resource under the decision option being considered
Morbidity	a) Number of people who suffer severe adverse health effects (illnesses or injuries) in a year attributable to the product or resource
	b) Number of people who suffer less severe adverse health effects (illnesses or injuries) in a year attributable to the product or resource
	c) Number of people who suffer other adverse health effects in a year attributable to the product or resource serious enough to affect quality of life
Personal controllability	Percentage of the exposed population who have sufficient knowledge and alternatives available that they could control or reduce their personal risk from the product being evaluated
Ability to detect adverse health effects	Any one of the following, as appropriate for the decision being evaluated: • Probability that a problem will be detected if it occurs • Frequency with which unexpected adverse effects would have to occur in the exposed population to be recognized • Increase in the base rate of occurrence of the adverse effect caused by the product being evaluated that would be sufficient to be recognized
Ability to mitigate adverse health effects	Probability that an informed institution will be able to reduce (mitigate) adverse health effects associated with the product being evaluated if such a problem is known to exist and is detected or alternatively, percentage of potentially affected people whose risks can be reduced (mitigated) by institutional actions

Using the Best Available Information to Develop Estimates

The data required to support the evaluation of the public-health consequences of different decisions can vary widely in quality and availability. For some problems, large volumes of epidemiologic data may be available to describe the mortality risk of some medical product. For others, there may be an array of detailed computer models that attempt to characterize the morbidity-reduction effectiveness of various inspection procedures. For still others, only scant data or models may be available, and it may be necessary to rely solely on the judgment of knowledgeable professionals to develop such estimates. Often, a hybrid approach will be necessary: using experts to identify relevant information from studies or models and relying on the experts to interpret the data or model results in the context of the decisions being evaluated. If data are available, their relevance to the required estimates must be considered carefully. For example, the available data may indicate deaths that were temporally associated with a product or device but were not caused by the product or device; judgment will be required to interpret such data appropriately and to estimate the mortality associated with the different decision options being evaluated.

BOX 2-2 What about Benefits?

In the committee's charge and interactions with FDA, FDA discussed the importance of characterizing the health benefits—in addition to the adverse health consequences—of the products that they regulate. Identifying product or product category benefits may be important in a pure risk-ranking activity as a way to recognize that some products may have favorable effects that should be balanced against their adverse effects before decisions are made. However, because the present framework focuses explicitly on comparing the outcomes of different decisions, the public-health benefit of one decision vs another is made clear by comparing the values of each attribute under the different decisions. For decisions that compare the outcomes of two or more decision options—such as mitigation-selection decisions to determine whether a drug should be removed from the market or strategic-investment decisions to determine whether postmarket surveillance of specific medical devices should be increased—positive and negative changes in the attributes can be used to capture the benefits of one decision vs another. For example, if a drug is removed from the market, additional illnesses may occur because of lack of medicine; if the drug is kept available, there could be adverse health outcomes associated with complications from the drug itself but fewer or less severe illnesses from the underlying cause that the medication is intended to treat. The benefit of the drug would be in the differences in the numbers and severity of illnesses. For that decision, evaluating both outcomes associated with the underlying disease and outcomes associated with the drug may facilitate the estimate of benefits. For targeting decisions, such as ranking various food groups by risk, the notion of benefits may not be relevant. If the goal of a study is to identify where fundamental risk exists (that is, where FDA should look to reduce adverse health effects), the study would focus on the adverse outcomes. If the question is which activity should have its budget cut to support increased surveillance elsewhere, the decision is similar to the one for removal of a drug from the market; benefits would be captured by comparing the outcomes of different levels of surveillance.

Critical decisions often must be made when uncertainty about some relevant factors remains; in those cases, judgment is required. Lack of easily accessible data does not mitigate the need to develop the best possible estimates of public-health consequences of products being evaluated; in some cases, carefully assessed expert judgment may be the only option available for ensuring that decision-makers have all the relevant information that is available to make decisions. Delaying a critical decision until additional information can be collected may be an option in some circumstances, but it should not be considered the "default" decision when uncertainty exists. Instead, all relevant options, including the option to delay a decision until later or to allocate scarce resources to additional data collection or model development, should be evaluated on the basis of the best available information, including expert judgment, about the public-health consequences of the decision.

Many risk-assessment and evaluation tools are available and are described in detail in other studies, including National Research Council and FDA studies described above. Box 2-3 summarizes current perspectives at EPA about how to quantify uncertain outcomes, including a hierarchy of approaches and the role of expert judgment. It is not within the present committee's scope to evaluate, compare, or recommend specific approaches or models for risk quantification. The committee simply urges FDA to bring the best available data and expertise to bear on the evaluation that would be consistent with the importance of the decision being evaluated and the time and resources available to complete the assessment.

Characterizing Uncertainty

Although substantial effort has been devoted to motivating and guiding the inclusion of uncertainty in policy analysis throughout the federal government (see Box 2-4), including it is difficult, and resistance remains high. Individual researchers and policy analysts are reluctant to make quantitative

estimates without what they feel to be "sufficient" data, and they are even more reluctant to quantify the uncertainties inherent in their domain. It is always possible to collect more data and do more analyses to try to develop "better" estimates, but there will always be uncertainty, and decisions often must be made on the basis of existing information. Quantifying what is known and what is not known (the uncertainty) is an important way to ensure that decisions are as well informed as possible.

The difficulty of developing best estimates and uncertainties when data are sparse is exacerbated by the norm of using probability distributions for quantification. Probability is not a natural way of thinking about uncertainty for everyone, and often efforts are made to sidestep the problem by using qualitative or categorical measures of uncertainty, such as "likely," "very unlikely," and "possible." That approach may make the risk-characterization task more palatable, but it also makes it considerably less useful. The very ambiguity that provides comfort makes the task of communicating and comparing

BOX 2-3 Methods for Estimating Uncertain Quantities

A white paper written for EPA outlines a general hierarchy of methods that can be used to estimate quantities when there is substantial uncertainty about their "true" values (Frey et al. 2003). Four general categories of methods are described without any implied preferences or priorities:

- Statistical methods based on empirical data, which use classical statistics to draw inferences from "hard" data alone.
- Statistical methods based on judgment, in which expert judgments and Bayesian approaches to statistical analysis are included, often in combination with "hard" data.
- Other quantitative methods that involve approaches not based on probability theory, such as interval methods, fuzzy methods, and meta-analytic methods.
- Qualitative methods that can be used when key aspects of uncertainty cannot be captured by quantitative methods (for example, uncertainty caused by problem formulation or the existence of competing models).

Of those approaches, the first is widely used and accepted in risk assessment, the second is widely used outside the risk-assessment community and is expanding in use and acceptance in the risk-assessment community, and the third and fourth approaches are not generally used.

EPA discussed the use of expert elicitation (or expert judgment) and made the following comment:

> Expert elicitation is recognized as a powerful and legitimate quantitative method for characterization of uncertainty and for providing probabilistic distributions to fill data gaps where additional research is not feasible. The academic and research community, as well as numerous review bodies, have recognized the limitation of empirical data for characterization of uncertainty and have acknowledged the potential for using [expert elicitation] for this purpose. In Science and Judgment in Risk Assessment [NRC 1994] the NAS [National Academy of Sciences] recognized that for "parameter uncertainty, enough objective probability data are available in some cases to permit estimation of the probability distribution. In other cases, subjective probabilities might be needed." In this "Blue Book" report, the NAS further recognized the "difficulties of using subjective probabilities in regulation" and identified perceived bias as one major impediment; but, noted that "in most problems real or perceived bias pervades EPA's current point-estimate approach." In addition, the NAS stated that "there can be no rule that objective probability estimates are always preferred to subjective estimates, or vice versa" (EPA 2009, p. 5).

Thus, EPA recognized the importance of using statistical methods based on judgment to derive probability estimates and emphasized that objective methods are not always preferable to subjective methods.

uncertainties difficult. Flexible definitions that vary from domain to domain make cross-cutting analysis impossible. Is the probability value associated with a "high likelihood" of death equal (either numerically or cognitively) to the probability value associated with a "high likelihood" of rain? The effort needed to define the uncertainty categories unambiguously and in sufficient detail would certainly approach if not exceed the effort needed to follow one of the standardized approaches of assessing a distribution directly. In addition, for some problem domains, assessing the complete distribution is not necessary; a few key points selected from the underlying distribution will describe the underlying uncertainty sufficiently to permit the necessary comparisons and analyses. That approach is demonstrated in the case studies that follow this chapter.

Options for succinctly describing a probability distribution vary; however, the shorthand notation should be able to provide an indication of a distribution's first three moments—central tendency, spread, and skew—and should be easily assessable from data, model output, or expert judgment. In the case studies that follow, that task is accomplished by using three values: the distribution's 5th, 50th (median), and 95th percentiles. The 5th percentile is the numerical value of the quantity being estimated (such as the number of deaths) that bounds the lower 5% tail of the distribution; for example, the probability that the actual value will be lower than the 5th percentile is 0.05. The 95th percentile is the numerical value that will be exceeded with a probability of 0.05; that is, it is the boundary of the upper 5% tail of the

BOX 2-4 Importance of Characterizing Uncertainty

The assessment, quantification, and communication of uncertainty are essential, as has been recognized in the risk-assessment literature since at least 1983. Numerous high-level government advisory panels have discussed the importance of capturing and presenting uncertain values. *Risk Assessment in the Federal Government: Managing the Process* (NRC 1983) formalized the integration of uncertainty and risk into policy analysis. The guidance provided in that early report has been expanded on in a series of reports (NRC 1993, 1994, 1996) and recently reinforced in *Science and Decisions: Advancing Risk Assessment* (NRC 2009, p. 93), which states at the beginning of a chapter on uncertainty and variability that "characterizing uncertainty and variability is key to the human health risk-assessment process, which must engage the best available science in the presence of uncertainties and difficult-to-characterize variability to inform risk-management decisions." In recent guidelines on probabilistic risk assessment, the U.S. Nuclear Regulatory Commission states that risk-informed decision-making requires the "appropriate consideration of uncertainty…in the analyses used to support the decision and in the interpretation of the findings of those analyses" (USNRC 2009, page iii).

Over the last 25 years, EPA has developed a series of guidelines that detail how uncertainty should be integrated into regulatory policy-making and evaluation (EPA 1984, 1995, 2000, 2004). *Estimating the Public Health Benefits of Proposed Air Pollution Regulations* (NRC 2002) offers advice to EPA on the importance of characterizing uncertainty even when data are sparse or lacking:

> EPA should move the assessment of uncertainty from its ancillary analyses into its primary analyses to provide a more realistic depiction of the overall degree of uncertainty. This shift will entail the development of probabilistic, multiple-source uncertainty models based not only on available data but also on expert judgment … Uncertainty should be described as completely and as realistically as possible for all regulatory options, recognizing that regulatory action might be necessary in the presence of substantial uncertainty. The regulatory decision process will be better informed by a fair assessment of the uncertainty and a realistic evaluation of the likely reductions in that uncertainty attainable through further research (NRC 2002, p. 11).

Explicit evaluation and presentation of uncertainty in risk assessments reduces the problem of false precision, makes risk characterization more informative, and increases the credibility of any ensuing risk communication.

distribution. The median value is a measure of the distribution's central tendency and generally represents the "best estimate" of the uncertain quantity. The probability that a value falls below the median is equal to the probability that it falls above the median. The spread of the distribution is indicated by the difference between the 5th and 95th percentiles. The skew of the distribution is indicated by comparing the difference between the 5th percentile and the median and the difference between the median and the 95th percentile. For a symmetric distribution, those differences will be equal. For a negative skew, the first difference will be larger; for a positive skew, the second difference will be larger.

If the three percentiles are relied on to characterize a distribution, the assessment procedure for experts is relatively simple and well tested. With training and practice, an analyst can elicit an accurate representation of the three percentiles from experts by using easily described thought experiments, simple tools, and standard protocols. Numerous references describe a variety of approaches for assessing quantitative estimates of uncertain quantities, including examples of how expert assessments have been used to support a wide array of decisions (Merkhofer 1987; Cooke 1991; Hora 2007; EPA 2009; Jenni and van Luik 2010). It is important to note that a viable approach will require the consideration of the tails of the distribution. Assessed distributions are often too narrow, but considering rarer and more extreme events leads to assessed distributions that represent true uncertainties better.

Although the three percentiles are precisely defined, their cognitive interpretation should not be lost. The goal is to have a risk characterization that includes representation of the uncertainty in the size of the exposed population, in the number of deaths that may occur, and in the number of injuries or illnesses that may occur under various decision options. Rather than ask simply for a range or for low and high estimates of those quantities, the committee suggests using the concepts associated with the percentiles of a distribution as a way to ensure consistent assessment and interpretation of the low and high values. The 5th percentile represents a value below which the actual value is not likely to fall. Similarly, the 95th percentile represents a value above which the actual value is not likely to fall. The estimates are not upper and lower bounds, but values that indicate limits beyond which results would be surprising. Fixating on the precise values for the percentiles could cause unnecessary anxiety and communicate false precision where none is intended or needed.

Uncertainty in the health consequences of alternative decisions should be included in the characterization of risks in this framework. Specifically, a best estimate, a high estimate, and a low estimate—corresponding to the three percentiles described above—should be determined for the size of the exposed population, mortality, and each of the three morbidity estimates. The other attributes could also be described by using similar distributional measures although the importance of doing so is less. When assessing the probabilities associated with ability to detect, control, or mitigate, ranges may be appropriate instead of point values. For example, the probability that an informed institution will be able to reduce or mitigate adverse health effects associated with a particular risk could be judged to be 0.80-0.90.

The committee notes that there is likely to be a relationship between the quality of data available to develop the necessary estimates and the uncertainty in the estimates. Specifically, when relevant data are lacking, uncertainty about each of the outcomes will probably be high (that is, a large spread).

USING THE FRAMEWORK TO SUPPORT DECISION-MAKING

The framework proposed here is similar to the framework for risk-based decision-making described in *Science and Decisions* (see Figure 2-1) in its emphasis on structuring and conducting risk assessments and risk characterization in the broader context of supporting agency decisions. *Science and Decisions* concludes by making recommendations for improving risk-based decision-making at EPA; although some of those recommendations may be applicable to FDA, the present committee was charged not with evaluating the risk-based decision-making processes at FDA but with developing a robust approach for characterizing public-health consequences. Those consequences are an important consideration in FDA decisions, and the attributes proposed in this framework provide such a

characterization. Using this framework throughout the agency will ultimately lead to robust and consistent characterization of public-health consequences.

How the information is used will ultimately be determined by FDA. Public-health consequences constitute only one of numerous factors—such as economic, social, and political factors—that FDA must consider when making decisions, and the committee neither expects nor recommends that the attributes in this report should form the sole basis of such decisions. However, careful and consistent evaluation of the public-health consequences of various options is an essential component of good decision-making, and there are several ways in which the risk characterizations described here could be used in decision-making.

One approach was outlined in the committee's letter report, which focused on developing a risk ranking (see Appendix A). The letter report describes an approach for involving stakeholders in an exercise to rank risks that used a consistent set of attributes. Specifically, it involved ranking risks on the basis of judgment and the described characteristics of those risks or ranking risks on the basis of a mathematical combination of the attribute scores and judgments about the importance of the different attributes, that is, using the logic of multiattribute utility theory (Keeney and Raiffa 1976). The approach described in the letter report is based on the risk-ranking literature described previously and could be adapted and applied in the decision-focused framework proposed here.

The risk characterization of Step 2 of the present framework is also compatible with several other approaches to decision-making—both more quantitative approaches and more inclusive processes. The analytic-deliberative process described in *Understanding Risk* (NRC 1996) includes scientists, public officials, and other interested and affected parties in an iterative process in which all parties are involved in every step of the decision process. If FDA adopted that type of approach to risk-based decision-making, the risk attributes defined here could provide a starting point for the risk-characterization aspects of the analysis. Retaining the flexibility to refine the risk characterization throughout the process, however, is key to the analytic-deliberative process.

At perhaps a different extreme, FDA may determine that for some decisions only a single public-health consequence is relevant. In that case, the other elements of the risk characterization could be ignored. Alternatively, the quantitative public-health consequences as a group may be considered relevant, and it may be considered useful to combine the attributes in the risk characterization into a single metric to compare options. Such integrative measures may be easier to develop, and especially valuable, when there is a logical set of mathematical relationships between the various attributes. For example, two rare causes of morbidity associated with a particular product might reasonably be considered additive in their effects; other attributes, such as the size of a population affected and the rate of adverse effects or the number of adverse effects and the likelihood that harm could be mitigated, would combine in a multiplicative fashion. Proven mathematical approaches for combining attributes into a single integrative measure (Keeney and Raiffa 1976; Fischhoff et al. 1984) could also be applied here. However, the attributes defined here were chosen to provide a robust risk characterization rather than an easy mathematical combination, so developing a single integrative measure would require careful consideration of the relationship between the various attributes and their relative importance. As discussed in Box 2-1, using integrative measures alone has several disadvantages, but relying only on the full list of attributes without integration can make consistency a challenge.

Each case study that follows concludes with a discussion of how the present framework could be used to support a specific decision, but they are offered only for illustration. It remains for FDA to decide how to use the information, and whether and how to combine it with other decision-relevant factors.

FLEXIBILITY AND EVOLUTION OF THE FRAMEWORK

The risk-characterization framework proposed here is a flexible system. Although it is intended to capture elements of risk that are applicable across the broad array of risk-related decisions that FDA makes, it is not necessarily a comprehensive set of public-health-related factors relevant for every such

decision, and it clearly does not aim to include factors that are unrelated to public-health outcomes. The framework is not proposed as a "one-size-fits-all" method for all risk-related decisions at FDA for all time. Just as the framework and the risk attributes were refined through development of the case studies in this report, the committee expects that some elements of the framework will continue to evolve as FDA gains experience in using it. Additional risk characteristics may be identified for subsets of decisions and added to the list for those types of decisions. Furthermore, the existing definitions of attributes may be modified to yield more interesting insights, and some attributes may be dropped altogether if they are not found to be useful.

REFERENCES

AHRQ (Agency for Healthcare Research and Quality). 2005. U.S. Valuation of the EuroQol EQ-5D Health States. Agency for Healthcare Research and Quality, Rockville, MD. December 2005 [online]. Available: http://www.ahrq.gov/rice/EQ5Dproj.htm [accessed Oct. 25, 2010].

Bertoni, M.J. 2010. Opening Remarks. Presentation at the 5th Meeting on Ranking FDA Product Categories Based on Health Consequences, Phase II, February 3, 2010, Washington, DC.

Bronfman, N.C., L.A. Cifuentes, M.L. DeKay, and H.H. Willis. 2007. Accounting for variation in the explanatory power of the psychometric paradigm: The effects of aggregation and focus. J. Risk Res. 10(4):527-554.

Bronfman, N.C., L.A. Cifuentes, and V.V. Gutierrez. 2008a. Participant-focused analysis: Explanatory power of the classic psychometric paradigm in risk perception. J. Risk Res. 11(6):735-753.

Bronfman, N.C., E.L. Vazquez, V.V. Gutierrez, and LA. Cifuentes. 2008b. Trust, acceptance and knowledge of technological and environmental hazards in Chile. J. Risk Res. 11(6):755-773.

Broome, J. 1993. QALYs. J. Public Econ. 50(2):149-167.

Cooke, R.M. 1991. Experts in Uncertainty: Opinion and Subjective Probability in Science. K. Shrader-Frechette, ed. New York: Oxford University Press.

Davies, J.C. 1996. Comparing Environmental Risks: Tools for Setting Government Priorities. Washington, DC: Resources for the Future.

DeKay, M.L., H.K. Florig, P.S. Fischbeck, M.G. Morgan, K.M. Morgan, B. Fischhoff, and K.E. Jenni. 2001. The use of public risk ranking in regulatory development. Pp. 208-230 in Improving Regulation: Cases in Environment, Health, and Safety, P.S. Fischbeck, and R.S. Farrow, eds. Washington, DC: Resources for the Future.

DeKay, M.L., P.S. Fischbeck, H.K. Florig, M.G. Morgan, K.M. Morgan, B. Fischoff, and K.E. Jenni. 2005. Judgement-Based Risk Ranking for Food Safety Pp. 198-226 in Toward Safer Food: Perspectives on Risk and Priority Setting. S.A Hoffman and M.R. Taylor, eds. Washington, DC: RFF Press.

EPA (U.S. Environmental Protection Agency). 1984. Risk Assessment and Management: Framework for Decision Making. EPA 600/9-85-002. U.S. Environmental Protection Agency, Washington, DC. December 1984.

EPA (U.S. Environmental Protection Agency). 1987. Unfinished Business: A Comparative Assessment of Environmental Problems. EPA/230/2-87/025. Office of Policy, Planning, and Evaluation, U.S. Environmental Protection Agency, Washington, DC.

EPA (U.S. Environmental Protection Agency). 1995. Policy for Risk Characterization at the U.S. Environmental Protection Agency. U.S. Environmental Protection Agency, Washington, DC. March 1995 [Appendix A in EPA 2000].

EPA (U.S. Environmental Protection Agency). 2000. Risk Characterization Handbook. EPA 100-B-00-002. Office of Science Policy, Office of Research and Development, U.S. Environmental Protection Agency. December 2000 [online]. Available: http://itp-pfoa.ce.cmu.edu/docs/RChandbk_excerpts_re_exposure.pdf [accessed Oct. 25, 2010].

EPA (U.S. Environmental Protection Agency). 2004. An Examination of EPA Risk Assessment Principles and Practices, Staff Paper Prepared for the U.S. Environmental Protection Agency by Members of the Risk Assessment Task Force. EPA/100/B-04/001. Office of the Science Advisor, U.S. Environmental Protection Agency, Washington, DC. March 2004 [online]. Available: http://www.epa.gov/osa/pdfs/ratf-final.pdf [accessed Oct. 25, 2010].

EPA (U.S. Environmental Protection Agency). 2009. Expert Elicitation Task Force White Paper, External Review Draft. Science Policy Council, U.S. Environmental Protection Agency. January 6, 2009 [online]. Available:

http://www.epa.gov/osa/pdfs/elicitation/Expert_Elicitation_White_Paper-January_06_2009.pdf [accessed Feb. 23, 2011].

EPA SAB (U.S. Environmental Protection Agency Science Advisory Board). 1990. Reducing Risk: Setting Priorities and Strategies for Environmental Protection. SAB-EC 90-021. Science Advisory Board, U.S. Environmental Protection Agency [online]. Available: http://yosemite.epa.gov/sab/sabproduct.nsf/28704D9C420FCBC1852573360053C692/$File/REDUCING+RISK++++++++++EC-90-021_90021_5-11-1995_204.pdf [accessed Mar. 17, 2011].

EuroQoL Group. 1990. EuroQol – A new facility for the measurement of health-related quality of life. Health Policy 16(3):199-208.

FDA/USDA (U.S. Food and Drug Administration and U.S. Department of Agriculture). 2003. Listeria monocytogenes Risk Assessment: Interpretive Summary. Center for Food and Applied Nutrition, U.S. Food and Drug Administration, and Food Safety and Inspection Service, U.S. Department of Agriculture. September 2003 [online]. Available: http://www.fda.gov/Food/ScienceResearch/ResearchAreas/RiskAssessmentSafetyAssessment/ucm185291.htm [accessed Feb. 23, 2010].

Fischhoff, B. 1984. Setting standards: A systematic approach to managing public health and safety risks. Manage. Sci. 30(7):823:843.

Fischhoff, B., S.R. Watson, C. Hope. 1984. Defining risk. Policy Sciences. 17(2):123-139.

Fischhoff, B. 1995. Ranking risks. Risk Health Saf. Environ. 6:189-200.

Fischhoff, B. 2006. Cognitive processes in stated preference methods. Pp. 937-968 in Handbook of Environmental Economics, Vol. 2. Valuing Environmental Changes, K.G. Mäler, and J.R. Vincent, eds. Amsterdam: Elsevier.

Florig, H.K., M.G. Morgan, K.M. Morgan, K.E. Jenni, B. Fischoff, P.S. Fischbeck, and M.L. DeKay. 2001. A deliberative method for ranking risks (I): Overview and test bed development. Risk Analysis. 21(5):913.

FoodRisk.org. 2011. FDA's Fresh Produce Risk Ranking Tool. Joint Institute for Food Safety and Applied Nutrition, College Park, MD [online]. Available: http://www.foodrisk.org/exclusives/RRT/ [accessed Feb.11, 2011].

Frey, C., D. Crawford-Brown, Z. Junyu, and D. Loughlin. 2003. Hierarchy of Methods to Characterize Uncertainty: State of Science of Methods for Describing and Quantifying Uncertainty. EPA Contract No. 68-D-00-265. Prepared for Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, NC. August 12, 2003 [online]. Available: http://www4.ncsu.edu/~frey/reports/Frey_CrawfordBrown_Zheng_Loughlin_2003.pdf [accessed Feb. 23, 2010].

Gutiérrez, V.V., L.A. Cifuentes, and N.C. Bronfman. 2006. The influence of information delivery on risk ranking by lay people. J. Risk Res. 9(6):641-655.

Hora, S.C. 2007. Eliciting probabilities from experts. Pp. 129-153 in Advances in Decision Analysis: From Foundations to Applications, W. Edwards, R.F. Miles, Jr., and D. von Winterfeldt, eds. New York: Cambridge University Press.

IOM (Institute of Medicine). 2006. Valuing Health for Regulatory Cost-Effectiveness Analysis. Washington, DC: The National Academies Press.

IOM (Institute of Medicine). 2009. Initial National Priorities for Comparative Effectiveness Research. Washington, DC: The National Academies Press.

Jenni, K.E. 1997. Attributes for Risk Evaluation. Ph.D. Dissertation, Carnegie Mellon University, Pittsburgh, PA.

Jenni, K.E., and A. van Luik. 2010. Assessment of expert judgments for safety analyses and performance assessment of geologic repository systems. Pp. 580-609 in Geological Repository Systems for Safe Disposal of Spent Nuclear Fuels and Radioactive Waste, J. Ahn, and M.J. Apted, eds. Cambridge: Woodhead Publishing.

Jones, K. 1997. A Retrospective on Ten Years of Comparative Risk. Prepared for American Industrial Health Council, Washington, DC, by Green Mountain Institute for Environmental Democracy, Montpelier, VT. January 24, 1997 [online]. Available: http://www.heartland.org/custom/semod_policybot/pdf/4728.pdf [accessed March 17, 2011].

Keeney, R.L. and H. Raiffa. 1976. Decisions with Multiple Objectives: Preferences and Trade-Offs. New York: Wiley.

Long, J. and B. Fischoff. 2000. Setting risk priorities: A formal model. Risk Analysis. 20(3):339-352.

Lopez, A.D., C.D. Mathers, M. Ezzati, D.T. Jamison, and C.J.L. Murray, eds. 2006. Global Burden of Disease and Risk Factors. Washington, DC: World Bank [online]. Available: http://www.dcp2.org/pubs/GBD [accessed Feb. 14, 2011].

Merkhofer, M.W. 1987. Quantifying judgmental uncertainty: Methodology, experiences, and insights. IEEE T. Syst. Man Cyb. 17(5):741-752.

Minard Jr., R.A. 1996. Can States Make a Market for Environmental Goals? Pp. 243-280 in Linking Science and Technology to Society's Environmental Goals. National Research Council. Washington, DC: The National Academies Press.

Morgan, M.G., B. Fischhoff, L. Lave, and P. Fischbeck. 1996. A proposal for ranking risks within federal agencies. Pp. 111-148 in Comparing Environmental Risks: Tools for Setting Government Priorities, J.C. Davies, ed. Washington, DC: Resources for the Future.

Morgan, K.M., M.L. DeKay, and P.S. Fischbeck. 1999. A multi-attribute approach to risk prioritization. Risk Policy Report 6(6):38-40.

Morgan, M.G., H.K. Florig, M.L. DeKay, and P. Fischbeck. 2000. Categorizing risks for risk ranking. Risk Analysis. 20(1):49-58.

Morgan, K.M., M.L. DeKay, P.S. Fischbeck, M.G. Morgan, B. Fischhoff, and H.K. Florig. 2001. A deliberative method for ranking risks (II): Evaluation of validity and agreement among managers. Risk Anal. 21(5):923-937.

NRC (National Research Council). 1983. Risk Assessment in the Federal Government: Managing the Process. Washington, DC: National Academy Press.

NRC (National Research Council). 1989. Improving Risk Communication. Washington, DC: National Academy Press.

NRC (National Research Council). 1993. Pesticides in the Diets of Infants and Children. Washington, DC: National Academy Press.

NRC (National Research Council). 1994. Science and Judgment in Risk Assessment. Washington, DC: National Academy Press.

NRC (National Research Council). 1996. Understanding Risk: Informing Decisions in a Democratic Society. Washington, DC: National Academy Press.

NRC (National Research Council). 2002. Estimating the Public Health Benefits of Proposed Air Pollution Regulations. Washington, DC: The National Academies Press.

NRC (National Research Council). 2009. Science and Decisions: Advancing Risk Assessment. Washington, DC: National Academy Press.

Slovic, P. 1987. Perception of risk. Science 236(4799):280-285.

Slovic, P. 1992. Perception of risk: Reflections on the psychometric paradigm. Pp. 117-152 in Social Theories of Risk, S. Krimsky, and D. Golding, eds. New York: Praeger.

USNRC (U.S. Nuclear Regulatory Commission). 2008. Risk-Informed Decisionmaking for Nuclear Material and Waste Applications: Revision 1. U.S. Nuclear Regulatory Commission, Washington, DC [online]. Available: http://www.hss.energy.gov/nuclearsafety/ns/rawg/docs/RIDM_2008_2.pdf [accessed Apr. 13, 2011].

USNRC (U.S. Nuclear Regulatory Commission). 2009. Guidance on the Treatment of Uncertainties Associated with PRAs in Risk-Informed Decision Making Main Report, Volume 1. NUREG-1855. U.S. Nuclear Regulatory Commission, Washington, DC [online]. Available: http://www.nrc.gov/reading-rm/doc-collections/nuregs/staff/sr1855/v1/sr1855v1.pdf [accessed Apr. 13, 2011].

Willis, H.H., M.L. DeKay, M.G. Morgan, H.K. Florig, and P.S. Fischbeck. 2004. Ecological risk ranking: Development and evaluation of a method for improving public participation in environmental decision making. Risk Analysis. 24(2):363-378.

Willis, H.H., M.L. DeKay, B. Fischoff, and M.G. Morgan. 2005. Aggregate, disaggregate, and hybrid analyses of ecological risk perceptions. Risk Analysis. 25(2):405-428.

3

Case Study of a Mitigation-Selection Decision

This chapter describes a detailed case study of the use of the risk-characterization framework to evaluate a hypothetical decision on whether to withdraw a vaccine from the market and provides information on the human-health consequences of two mitigation options that could be used as input in a comprehensive decision-making process. The case study described was selected because it involves a mitigation-selection decision, it is related to several decision scenarios provided to the committee by the Food and Drug Administration (FDA), and several committee members had knowledge of the case history. The data used were gleaned from publicly available Web sites or publications or provided by FDA. The committee did not conduct exhaustive literature searches or reviews, and all information is illustrative. It emphasizes that the case study simply provides an illustration of how the committee's framework might be used for a mitigation-selection decision.

FRAMING THE ISSUE: VACCINE WITHDRAWAL

Before vaccine availability, the burden of rotavirus infection in the United States was estimated to be 3.5 million cases per year in children less than 5 years old (Murphy et al. 2001). Rotavirus caused gastroenteritis in nearly all infected children and serious complications in a small fraction of that number. Using several previously published studies, Glass et al. (1996) estimated that the annual number of hospitalizations for rotavirus diarrhea in the United States ranged from 23,000 to 110,000; the best projected estimate was 55,000 hospitalizations per year. They estimated the number of deaths from rotavirus diarrhea in the United States at 20-40 per year. It was also projected that physician visits were less frequent than cases of diarrhea.

In three randomized prelicensure trials, the first available tetravalent rhesus-human reassortant rotavirus vaccine, RRV-TV (RotaShield, Wyeth Lederle Vaccines, Philadelphia), was shown to be highly effective (80-100%) for the prevention of severe rotavirus gastroenteritis in infants (Rennels et al. 1996; Perez-Schael et al. 1997; Santosham et al. 1997; Joensuu et al. 1998).[1] In the conduct of 27 prelicensure trials of RRV-TV, five cases of intussusception—a rare form of bowel obstruction in which a portion of the bowel prolapses into a more distal portion—were reported in 10,054 infants who received the vaccine compared with only one case in 4,633 recipients of placebo; the rates were not statistically significantly different ($p > 0.45$) (Rennels et al. 1998). After much deliberation about the potential risk of intussusception after RRV-TV, FDA approved RRV-TV on August 31, 1998, for administration at the ages of 2, 4, and 6 months. Intussusception was listed as a possible adverse reaction in the manufacturer's product insert and in the published recommendations of the Advisory Committee on Immunization

[1]Severe rotavirus gastroenteritis is defined as diarrhea with or without vomiting that requires hospitalization or rehydration in a medical facility; stool-sample analysis confirms rotavirus infection (other than vaccine strain) (CDC 2009).

Practices (ACIP) of the Centers for Disease Control and Prevention (CDC) and the American Academy of Pediatrics (AAP 1998; CDC 1999a).

From September 1, 1998, to July 7, 1999, 15 cases of intussusception in infants who had received RRV-TV were reported to the Vaccine Adverse Event Reporting System (VAERS). There were no data on how many infants received RRV-TV; the manufacturer estimated that it had distributed 1.8 million doses and that about 1.5 million doses had been administered as of June 1, 1999 (CDC 1999b). Although some studies indicated that that number of cases would be expected by chance alone (CDC 1999b), the facts that vaccine-induced intussusception had been identified as a potential issue in prelicensure trials and that adverse events after vaccination are typically under-reported to the VAERS raised concerns that the actual number of intussusception cases in RRV-TV recipients might be substantially greater than the 15 reported. That concern led to the implementation of a multistate investigation to determine whether an association existed between administration of RRV-TV and intussusception in infants. Data from a postlicensure study at Northern California Kaiser Permanente were also quickly reviewed. The resulting data from both studies suggested that there was an increased risk of intussusception after receipt of RRV-TV. On July 16, 1999, CDC recommended suspension of the RRV-TV immunization programs until additional data could be analyzed (CDC 1999b; Alpert 1999).

After review of all available data, Wyeth voluntarily withdrew RRV-TV from the market on October 21, 1999. The next day, the ACIP withdrew its recommendation of infant vaccination with RRV-TV on the basis of the increased rate of intussusception after its administration (CDC 1999c). Considerable discussion followed the withdrawal of the RRV-TV vaccine and the implications for future vaccine development and evaluation in developed and developing countries, but that discussion is outside the scope of this exercise.

DECISION CONTEXT FOR THE CASE STUDY

The committee chose this real-world case to illustrate the use of the risk attributes defined in this report to organize, evaluate, and compare the public-health consequences associated with a mitigation-selection decision similar to others that FDA often faces. To complete the case study, the committee attempted to look at the situation and the decision options as they were understood in late summer to early fall 1999. Although Wyeth withdrew the vaccine from the market at that point, the committee assumed for purposes of the case study that FDA was in a position in which it had to make a decision about vaccine withdrawal.

The committee considered two options: leave the vaccine on the market or withdraw it from the market. Two types of health consequences are particularly relevant for this comparison. First, adverse effects after vaccination would affect all infants who were vaccinated and would be relevant only if it were decided to leave the vaccine on the market. Second, although the likelihood of contracting rotavirus disease is much lower with vaccination, adverse effects from rotavirus disease would be relevant for both options. Both types of effects would have to be considered for FDA to make a fully informed decision between the two options, so both are included as key parts of the risk characterization.

CHARACTERIZING THE PUBLIC-HEALTH CONSEQUENCES

The committee used data that were available in late 1999 (the hypothesized time of this decision) to develop estimates of the human-health consequences of the two options. Although data were available, as indicated in the detailed development of the estimates below, characterizing the public-health consequences of the two decision options required the committee to make informed judgments, some assumptions, and a variety of calculations. For example, estimates of the number or percentage of children who would receive the vaccine were required. Although substantial data on vaccination rates in the United States are available from CDC, the committee had to make judgments about the relevance of

various vaccinations rates to the rate of RRV-TV vaccination. Furthermore, developing estimates of the number of children who would suffer adverse health effects of vaccination required combining estimates of the number of children vaccinated and the adverse-effect rate for the vaccine, both of which are uncertain.

Table 3-1 summarizes the results of the risk characterization. The bases of the estimates, including assumptions and intermediate calculations, are discussed below.

Exposed Population

By the age of 5 years, almost everyone has been infected with rotavirus, and if people contract it after the age of 5 years, they generally do not suffer the complications of severe dehydration that are much more common in younger children. All children under 5 years old are considered to be at risk for symptomatic rotavirus infection and are identified as part of the exposed population for this case study. The susceptible population was estimated to be 20 million (assuming a birth cohort of 4 million each year) if RRV-TV were removed from the market. Longitudinal studies and placebo-controlled field trials of rotavirus vaccines that used active surveillance methods, including home visits twice each week, indicated that the incidence of rotavirus diarrhea reached 0.30 episodes per child per year during the first 2 years of life with a cumulative incidence approaching 0.80 episode per child by the age of 5 years (Glass et al. 1996). Those estimates suggest that about 75% of all rotavirus cases each year occur in children under 2 years old. Older children and adults can be reinfected with rotavirus, but the clinical symptoms are generally much less severe because previous infections confer some degree of immunity.

With the vaccine on the market, two populations were of interest: those exposed to the risks of adverse effects after RRV-TV vaccination (all infants who received the vaccine) and those exposed to the risks of rotavirus disease (all children under 5 years old who did not receive the RRV-TV and those who did receive the vaccine but who remained susceptible). Estimating the size of the two populations required estimates of vaccination rates and the vaccine effectiveness.

Because RRV-TV was not on the market for very long before it was withdrawn, the committee did not think it was appropriate to use the vaccination rates associated with it to calculate the vaccination rates that would have been accomplished by a fully mature program. Instead, data on overall vaccination rates in the United States for vaccines available at that time were reviewed. Using publicly available data from CDC, the committee examined the national average rates of vaccination for vaccines that were recommended in 1998 for children up to 24 months old. The rates ranged from 38% for varicella vaccination to 95% for three or more doses of DTP vaccine (CDC 2001). The committee concluded that the ultimate acceptance and adoption of RRV-TV would probably be less than those of DTP vaccines and decided to use the 95% vaccination rate as an estimate of the highest rate of vaccination with RRV-TV, representing about the 95th percentile of the distribution. Varicella vaccination rate was used as the estimate of the lowest rate of vaccination with RRV-TV, representing the 5th percentile of the distribution.

The committee notes that there are two possible reasons for the varicella vaccination rate to have been so much lower than the DTP rate, each of which is relevant for RRV-TV. First, the vaccine was approved and recommended for use in the United States in 1995; by 1998, it would have been offered to all children under 2 years old but would still be relatively novel in the market. Second, varicella is a common disease that is viewed as manageable by parents, so they may be less inclined to vaccinate. As a median estimate of the rate of RRV-TV use, the committee used 75% on the basis of the rates of having received a full complement of recommended vaccines in 24-month-olds in 1998. As noted above, estimating the rates required informed judgments to be made about the relevant data and their applicability to the future vaccination rates with RRV-TV.

TABLE 3-1 Risk Attributes for Mitigation-Selection Decision for Biologics

		Decision option	
Attribute	Metric	RRV-TV remains on market	RRV-TV is removed from market
Exposed population	Number vaccinated annually	3 million (1.5-3.8 million)	0
	Number who are susceptible to rotavirus disease annually	6.5 million (2.9-13.2 million)	20 million
	Populations of concern	Infants are vaccinated at 2, 4, and 6 months old; susceptible population for rotavirus disease in the unvaccinated population is the same as if vaccine were removed from market	All children under 5 years old are at risk for rotavirus disease; children under 2 years old are most at risk (~75% of cases) and experience more severe effects
Mortality	Number of deaths per year	15 (6.5-24) total: 9.8 (2.9-21) from rotavirus, 5.5 (0-10) from vaccine-induced intussusception	30 (20-40) from severe rotavirus disease
Morbidity	Number experiencing severe adverse health effects per year	18,300 (5,800-44,200) total: 17,900 (5,500-43,900) from rotavirus, 420 (0-620) from vaccine-induced intussusception	55,000 (23,000-110,000)
	Number experiencing less severe adverse health effects per year	162,500 (52,800-348,000) from rotavirus	500,000 (365,000-672,000)
	Number per year experiencing adverse health effects that affect only quality of life	1 million (375,000 to 2.05 million) total: 960,000 (324,000 to 2 million) from rotavirus, 42,000 (0-60,000) from post-vaccine fever or diarrhea	2.9 million (2.2-3.5 million)
Personal controllability	For vaccine-induced adverse effects	Parents of 80-90% of vaccine-eligible children have the ability to avoid or reduce the risks to their child	Not applicable
	For rotavirus disease	Not controllable	Not controllable
Ability to detect adverse health effects	Ability of informed institution to detect population-level effects associated with product being evaluated	Adverse effects would have to occur at a rate at least 5 times higher than normally expected to be detected; very-long-latency adverse effects are unlikely to be detected	Not applicable
Ability to mitigate adverse health effects	Probability that an informed institution will be able to reduce or mitigate adverse health effects associated with the product being evaluated if such a problem is known to exist	For RRV-TV induced effects: >99%	Not applicable

Vaccination rate and vaccine effectiveness, with the sizes of the birth cohort and the exposed population, are sufficient to develop estimates of the exposed populations of interest: those exposed to the risk of vaccine-associated side effects and those exposed to the risks of rotavirus disease. All children who receive RRV-TV will be exposed to the risk of adverse effects associated with the vaccine, including intussusception; the size of the exposed population is simply the birth cohort of 4 million multiplied by the vaccination rate. The population exposed to the risk of adverse effects after vaccination is therefore 3 million, with a range of 1.5 million to 3.8 million.

The children who would be susceptible to rotavirus disease if it were decided to leave the vaccine on the market include those who were not vaccinated and those who received the vaccine but for whom the vaccine is not effective. As stated above, RRV-TV was estimated to be 80-100% effective in eliminating severe rotavirus disease. For the purposes of this case study only, the committee applied the estimates of vaccine effectiveness (80-100% with a median of 90%) to all incidences of rotavirus disease rather than only to that of severe rotavirus disease. That simplification likely underestimates the size of the exposed population with the vaccine on the market and underestimates the number of less severe adverse effects (discussed in the next section). Thus, the number susceptible to rotavirus disease can be estimated from the following equation:

number susceptible to rotavirus = (20 million) [(1 - VR) + (VR) (1 - VE)],
where VR = vaccination rate and VE = vaccine effectiveness.

That equation can be used to calculate the number susceptible to rotavirus disease for any given estimates of vaccination rate and vaccine effectiveness. However, using the estimates to derive the 5th and 95th percentiles of the number of susceptible children requires some additional calculations; it is not as simple as using the 5th and 95th percentiles of the estimated vaccination rate and vaccine effectiveness. One approach is to use an event tree to evaluate all the possible combinations of vaccination rate and vaccine effectiveness and then to calculate the probabilities of the resulting estimates of susceptible children to identify which estimates are the 5th, 50th, and 95th percentiles. Figure 3-1 illustrates the event tree; the number of susceptible children under 5 years old is shown on the right side of the figure for every combination of vaccination rate and vaccine effectiveness considered. For example, given a vaccination rate of 38% and vaccine effectiveness of 80%, the number of susceptible children under 5 years old is 13,920,000. Using standard decision-analysis methods to assign a probability to each branch of the event tree (Keefer and Bodily 1983; Clemen 1996), the committee calculated a probability associated with each outcome and selected the 5th, 50th, and 95th percentiles of that combined distribution (highlighted) to represent the size of the population susceptible to rotavirus disease in the scenario in which RRV-TV is left on the market.

Table 3-1 summarizes the estimates of the number of infants at risk for vaccine-induced adverse effects and those at risk for rotavirus disease for the mitigation-decision options being compared.

Mortality and Morbidity

To estimate mortality and morbidity for the two decision options, the committee developed rough estimates of the following factors on the basis of publicly available data:

- The number of infants vaccinated each year (see discussion above).
- The effectiveness of the vaccine in protecting infants from rotavirus infection (see discussion above).

- The rotavirus infection rate in the unprotected population.
- The direct effects of rotavirus disease, characterized in terms of the rates of death and severe, less severe, and adverse quality-of-life health effects given rotavirus infection.
- The rate of adverse effects attributable to RRV-TV.
- The adverse effects attributable to RRV-TV, characterized in terms of rates of death and severe, less severe, and adverse quality-of-life health effects.

Rotavirus affects most infants in the United States but causes severe gastroenteritis in only a small fraction of those infected. As noted earlier, before the introduction of a vaccine, there were an estimated 3.5 million cases per year in children less than 5 years old in the United States, which led to 500,000 office visits ("less adverse events"), 55,000 hospitalizations ("severe adverse events"), and about 30 deaths (Glass et al. 1996; Fischer et al. 2007)—none of those numbers is known with precision. Table 3-2 summarizes the estimates of the number of rotavirus cases and the morbidity and mortality from rotavirus developed by the committee. The lower half of the table shows incidence rates calculated directly from the median estimates in the upper half of the table; footnotes to the table document the basis of the estimates. The incidence rates were used to calculate the numbers of deaths and other adverse health effects in various vaccination scenarios.

The committee recognized that the risk of serious rotavirus disease was much higher in some countries other than the United States (Dennehy 2008). However, because the focus of this study is on U.S. users of FDA-regulated products, the committee relied on and presents only estimates for the U.S. population, although it understands that other public-health agencies may be interested in health consequences that are outside the scope of FDA's authority to consider.

% vaccinated	Vaccine effectiveness	No. not protected annually	No. unprotected under 5 years old
	80%	2,784,000	13,920,000
38%	90%	2,632,000	13,160,000
	100%	2,480,000	12,400,000
	80%	1,600,000	8,000,000
75%	90%	1,300,000	6,500,000
	100%	1,000,000	5,000,000
	80%	960,000	4,800,000
95%	90%	580,000	2,900,000
	100%	200,000	1,000,000

FIGURE 3-1 Estimating the size of the unprotected population given uncertainty in vaccination rates and vaccine effectiveness. Highlighted cells show the 5th, 50th, and 95th percentiles of the combined distribution.

TABLE 3-2 Estimated Numbers and Rates of Deaths and Illnesses from Rotavirus Disease in the Absence of Vaccination

	Low Estimate (5th percentile)	Median Estimate	High Estimate (95th percentile)
Number of children			
Susceptible population[a]		20,000,000	
Total cases	2,800,000[c]	3,500,000[b]	4,000,000[d]
Deaths[b]	20	30	40
Severe adverse health effects[b]	23,000	55,000	110,000
Less severe adverse health effects	365,000[e]	500,000[b]	672,000[e]
Adverse health effects affecting only quality of life[f]		2,944,970	
Rates among susceptible children			
Deaths	0.000001	0.0000015	0.000002
Severe adverse health effects	0.00115	0.00275	0.0055
Less severe adverse health effects	0.018	0.025	0.034
Adverse health effects affecting only quality of life		0.147	

[a]Uncertainty in the number of children under 5 years old is insignificant relative to other uncertainties, so only a best estimate is used.
[b]Based on Glass et al. (1996).
[c]Based on Tucker et al. (1998), with an estimate of a 70% cumulative incidence by the age of 5 years.
[d]Assumes that all children under 5 years old get rotavirus disease; some cases are sufficiently mild that they might not be recognized as rotavirus disease.
[e]Based on assumptions described in Tucker et al. (1998). Low estimate based on assumption that 10% of all physician, clinic, and emergency-room visits for diarrhea are due to rotavirus; high estimate based on assumption that 25% of physician and clinic visits and 30% of emergency-room visits for diarrhea are due to rotavirus.
[f]Calculated value: all cases that do not lead to the more serious effects are assumed to have an adverse effect only on quality of life. 5th and 95th percentiles, not shown in this table, are calculated by using the event-tree approach described above and assuming independence between the rates for more serious effects.

RRV-TV was administered to nearly 7,000 infants in placebo-controlled studies and to 4,740 infants in three non-placebo-controlled trials before licensure (CDC 1999a). In those studies, 2% of all vaccine recipients experienced fever greater than 102.2ºF compared with 1% of all placebo recipients. In the placebo-controlled trials, investigators found no overall difference in the rate of diarrhea; but in studies in Finland, vaccinated children had a significantly higher rate of diarrhea after the first dose of vaccine than did placebo recipients (2.8% vs 1.4%; $p < 0.05$) (CDC 1999a). Because the adverse effects associated with vaccine administration were well tolerated (Joensuu et al. 1997) and generally did not require additional physician or emergency-department visits, the committee characterized them as adverse health consequences that affect only quality of life, similar to those cases of symptomatic rotavirus disease that do not require medical intervention. A 1.4% rate of postvaccine adverse effects was used as the median estimate of the rate of adverse effects affecting only quality of life, with a range of 0-2%.

At the time of the decision concerning withdrawal of the vaccine from the market, there was substantial uncertainty about the quantitative relationship between RRV-TV and increased rates of intussusception. A large case-control study that included 19 states where 80% of the RRV-TV vaccine was administered confirmed the association and was published in the *New England Journal of Medicine* in February 2001. The report concluded that there was an increased risk of intussusception 3-14 days after

both the first and second doses of RRV-TV but that the risk after the second dose was less than that after the first dose. With an assumption of full implementation of a national program of vaccination with RRV-TV, it was estimated that one case of intussusception attributable to the vaccine would occur in every 4,670-9,474 infants vaccinated (Murphy et al. 2001). The committee chose the midpoint of that range as the estimate of the median value, that is, about 1 in 7,000. It would have been reasonable to consider the possibility that there was no RRV-TV-attributable increase in the rate of intussusceptions because prelicensure studies had not established the association, and the 15 cases reported after RRV-TV vaccination were consistent with what would have been expected from base rates in the population (CDC 1999b). On that basis, the committee included a rate of RRV-TV-attributable intussusception of 0 as the estimate of the 5th percentile and a rate of 1 in 4,670 as the estimate of the 95th percentile.

The final estimates required are those of mortality and morbidity rates associated with intussusception. The mortality rate associated with intussusceptions has recently been estimated at about 1.3% (Cortese et al. 2009), which is consistent with data that would have been available in 1999, and was chosen as the median estimate for this case study. Parashar et al. (2000) used data from 1994-1997 to provide a range of estimates of the incidence of intussusception hospitalization of infants; the estimates provided were 18-56 cases per 100,000 children (1994-1996) and about 2.3 intussusception-caused deaths per million live births (1995-1997). Using those estimates, the committee projected the 5th and 95th percentiles of the distribution on the mortality rate with intussusception to be 0.4% and 1.9%, respectively. For purposes of this case study, the committee assumed that all occurrences of intussusceptions led to hospitalization (severe adverse health effects), and thus the rate of severe adverse effects of RRV-TV was assumed to be the same as the rate of vaccine-induced intussusception.

Estimates of Deaths and Morbidity for Option to Remove RRV-TV from the Market

The health consequences of a decision to withdraw the vaccine from the market can be estimated directly from the number of cases that occurred per year before the vaccine was available and the severity of the disease effects as described above and summarized in Table 3-2.

Estimates of Deaths and Morbidity for Option to Leave RRV-TV on the Market

Estimating the health consequences of a decision to keep the vaccine on the market requires two sets of estimates: one for the effects of rotavirus disease in susceptible people who did not receive vaccine or people who were vaccinated but in whom the vaccine was not effective and the other for the adverse effects associated with the vaccine itself. For purposes of this case study, the committee applied the mortality and morbidity rates for rotavirus disease estimated from the rates in the general population before 1998 to the unprotected population under different vaccination assumptions. That approach is a substantial simplification of what might be expected to occur. For example, widespread use of the vaccine could reduce the overall incidence in the unprotected population simply because prevalence of the circulating virus is reduced. That concept of vaccine protection is called herd immunity and probably would be present in a fully mature rotavirus-vaccination program. Similarly, the vaccine itself could be more effective in reducing the most severe cases of the disease and less effective in reducing the overall incidence. Again, for purposes of this case study, the committee assumed a single estimate of vaccine effectiveness and used it to reduce the size of the susceptible population rather than adjusting the incidence for the degree of severity of health effects in the vaccinated.

To estimate the numbers of deaths and severe adverse health effects attributable to vaccination, the committee used the rates of intussusception after RRV-TV and the mortality rate from intussusception described above and applied them to the size of the vaccinated population. The number of deaths for this decision option was estimated as follows:

$$N_f = R_f + I_f \quad \text{(Eq. 3-1a)}$$
$$R_f = [P_{<5} \times (1 - VR) + P_{<5} \times VR * (1 - VE)] \times M_R \quad \text{(Eq. 3-1b)}$$
$$I_f = P_{<1} \times VR \times I|V \times M_I , \quad \text{(Eq. 3-1c)}$$

where

N_f = total number of deaths, R_f = deaths from rotavirus disease, and I_f = deaths from vaccine-induced intussusception,
$P_{<5}$ and $P_{<1}$ represent the total population under the ages of 5 years and 1 year, respectively,
VR = vaccination rate,
VE = vaccine effectiveness,
I|V = rate of RRV-TV-induced intussusception, and
M_R and M_I represent mortality rates from rotavirus disease and intussusception, respectively.

Four of the factors (VR, VE, I|V, and M_R) are considered sufficiently uncertain that a range of estimates was developed and are described above. The relationships of the various factors are illustrated graphically in Figure 3-2.

Table 3-3 uses Equations 3-1(a-c) and median estimates for each of the factors documented above to illustrate a sample calculation of the number of deaths from RRV-TV-induced intussusception and from rotavirus disease with a decision to leave the vaccine on the market. Because the vaccination rate affects the number of deaths from rotavirus disease and the number of deaths from vaccine-induced intussusception in opposite directions (that is, the more children who are vaccinated, the fewer the deaths from rotavirus disease and the more deaths from vaccine-induced intussusception), estimating the range for the total number of deaths (or other adverse health effects) is not as simple as adding the low (5th percentile) or high (95th percentile) estimates of the number from each cause; doing so would overestimate the range. The committee used the decision-analysis approach described previously to propagate uncertainties in each factor and derive the final calculated value of the total number of deaths. The 5th, 50th, and 95th percentiles of that distribution are summarized in Table 3-1.

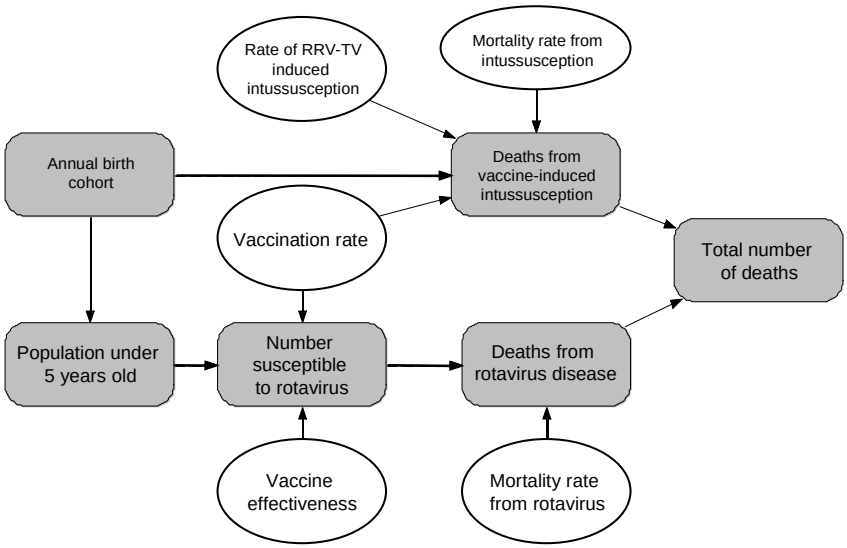

FIGURE 3-2 Relationship among various factors used to estimate the number of deaths from rotavirus disease and adverse effects of vaccination for the hypothetical decision to leave RRV-TV on the market. Quantities for which estimates of ranges of values were available are shown with ovals, and calculated values or quantities for which only a single estimate was available are shown as rounded rectangles.

Simply substituting the rates of severe adverse health effects and less severe adverse health effects for the mortality rates in Equations 3-1b and 3-1c yields estimates of the total number of children who would experience those effects if the vaccine were retained (see example calculation in Table 3-3). Those estimates are also summarized in Table 3-1.

TABLE 3-3 Example Calculation of Estimated Number of Deaths and Illnesses from Rotavirus Disease and RRV-TV-Induced Intussusceptions for Option to Leave the Vaccine on Market (Based on Median Estimates of all Factors)

Exposure estimates	
No. of children under age of 5 years	20,000,000
Size of annual birth cohort	4,000,000
Vaccine effectiveness	90%
Vaccination rate	75%
No. vaccinated per year	3,000,000
No. of susceptible children	6,500,000
Rotavirus infection and rate of adverse effects in the susceptible population[a]	
Rotavirus infection rate	17.5%
Mortality rate	0.00015%
Severe-disease rate	0.28%
Less-severe-disease rate	2.5%
Rate of only quality-of-life effects	14.7%
No. of deaths	9.8
No. of severe adverse effects	17,875
No. of less severe adverse effects	162,500
No. of only quality-of-life effects	957,115
Adverse health effects from RRV-TV vaccination	
Vaccine-induced intussusception rate	0.014%
Mortality rate of intussusception	1.3%
Rate of severe adverse effects of intussusception[b]	98.7%
Rate of only quality-of-life effects of vaccination (postvaccination fever and diarrhea)	1.4%
No. of deaths from RRV-TV attributable to intussusception	5.5
No. experiencing severe adverse health effects from RRV-TV-induced intussusception	423
No. experiencing only quality-of-life effects (from vaccination side effects)	42,000
Totals	
No. of deaths	15
No. of severe adverse health effects (hospitalizations)	18,298
No. of less severe adverse health effects (doctor visits)	162,500
No. of only quality-of-life effects	999,115

[a]Rates are calculated by dividing the median estimate of the number of cases of each type of health effect by the size of the susceptible population of 20 million; rates shown in the table are rounded to two or three significant figures, but the actual rates are used in the calculations. Differences in the number of effects shown in this table from the results of hand calculations using the rates shown in this table are due to rounding of the rates in the table.
[b]All incidents of intussusception that do not lead to death result in a "severe adverse health effect."

Personal Controllability

The primary way in which people can eliminate or reduce the RRV-TV-induced risks to their children is by declining to vaccinate, which would eliminate the risk of vaccine-induced illnesses. Controlling or reducing risks after vaccination is more difficult. Vaccine information sheets are provided to all parents before children are vaccinated and should make them aware of potential risks, normal side effects, and what adverse side effects warrant medical attention. Although those factors suggest that the personal controllability of vaccine-induced risks is high in theory, it can be argued that vaccinations are not perceived as being voluntary. For some vaccines, refusal to vaccinate may carry important consequences that make it impractical not to do so, such as the inability to attend day care or public school or to enlist in the armed forces.

The committee considered the varicella vaccine to be a reasonable analogue to the rotavirus vaccine, for reasons discussed above. When the varicella vaccine was first recommended in 1997, only about 26% of children received the vaccine. Immunization rates have increased substantially, but the relatively low early rates likely reflect a reluctance of physicians to recommend the vaccine soon after the universal recommendations and the choice of parents to decline the vaccine at that time. Of the 26% of parents who chose to have their children vaccinated, however, it is reasonable to assume that most considered the benefits of the vaccine to be worth the known risks posed by vaccination rather than feeling that they had no choice in the matter. Although the vast majority of parents choose to have their children vaccinated with recommended vaccines, the committee estimates that 80-90% of parents understand that they have the right to decline vaccinations and to control the RRV-TV-induced risks if they so desire.

In the absence of the vaccine, the ability of parents to prevent their child from contracting rotavirus disease is virtually zero. Rotavirus affects almost all children at least once during the first 5 years of life in developed and developing countries alike (Bernstein 2009).

Ability to Detect Adverse Health Effects

Some adverse effects after vaccination are expected; side effects that were identified during the testing and licensing process would be expected to occur at about the same rate in the vaccinated population. Of interest here are adverse effects that are not expected or that occur at a much higher rate than expected; such effects are an indication of a problem with the vaccine that could be of concern. In the case of vaccines, reporting systems are in place to track adverse effects after vaccination on a regular basis, such as VAERS and the CDC Vaccine Safety Datalink (VSD). In the near future, the Post-licensure Rapid Immunization Safety Monitoring (PRISM) system, which is part of FDA's Sentinel Initiative, will also contribute information. Those systems are intended to identify or refine signals of adverse effects that occur with increased frequency after vaccination.

In the case of RRV-TV, VAERS reporting and the fact that vaccine-induced intussusceptions had been identified as a potential issue led to rapid identification of the signal, although reports of intussusceptions were about equal to the expected size of the effect. It is widely recognized, however, that passive surveillance systems, such as VAERS, have issues with under-reporting. For example, one study (Braun 2006) estimated that the proportion of adverse events reported to VAERS after vaccination was 4% for one vaccine and 68% for another; the findings demonstrated that under-reporting can vary greatly depending on the type of adverse event and other factors. So the number of intussusceptions reports after RRV-TV vaccination was interpreted as implying a higher rate in practice. Additional vaccine-safety systems, such as the CDC VSD, provided data to establish more definitive assessments of the risk (Kramarz et al. 2001). More advanced vaccine-safety monitoring programs are in place now[2]

[2]Reconstruction of the VSD data on rotavirus vaccine suggests that it would have identified this problem if the current program had been active at that time (Lieu et al. 2007).

than in 1998, when RRV-TV was licensed; however, the ability to detect near-term adverse effects that were occurring at rates higher than expected was high even then. Using the metrics proposed in Chapter 2 and considering the many difficulties associated with interpreting data from passive surveillance (VAERS) (Ellenberg and Chen 1997), the committee estimated that in 1998-1999, the rate of an adverse effect of a vaccine would have to be at least 5 times the expected rate for it to be rapidly detected through the surveillance systems in operation at that time.

For any drug or vaccine, there is the possibility of unexpected adverse effects—effects too rare to be detected in prelicensure testing. If such effects arose with RRV-TV on the market and the effects were serious (ones that led to death or severe adverse health effects) and occurred at a rate substantially higher than would be expected for this population, the reporting systems in place at that time would have identified those effects quickly. The committee estimates that serious unexpected adverse effects from RRV-TV that occurred at a rate at least 5 times higher than normally occurs in the population would be detected. Detecting unexpected vaccine-induced adverse health outcomes that have a long latency, such as adverse effects that occur many years after vaccination, would be much more difficult. The committee estimates that the ability to detect such effects and determine that they are associated with a specific vaccine (RRV-TV) would be virtually zero.

The committee notes that at the time of this hypothetical evaluation, RRV-TV had been in use for less than a year. If it were removed from the market, the only possible adverse effect would be the loss of protection against rotavirus disease provided by the vaccine, and that is captured in the mortality and morbidity estimates for the vaccine-withdrawal option. Thus, there would be no other adverse effects of the absence of the vaccine to be detected, so this attribute does not apply to the decision option to withdraw the vaccine.

Ability to Mitigate Adverse Health Effects

As described in Chapter 2, this attribute refers to the ability of institutions to manage, reduce, or otherwise control adverse effects of the product being evaluated. Mitigation of adverse effects is relevant for effects that are expected and for effects that are not. For example, the routine use of an antipyretic before vaccination reduces the likelihood of fever after vaccination, a common and expected side effect. For adverse effects that occur at rates higher than expected or acceptable, other mitigation steps might be necessary. Because vaccines are given in the controlled setting of clinics and health-care centers, immediate adverse effects—such as anaphylaxis and hypersensitivity reactions—can be quickly identified and managed. Similarly, vaccine-induced intussusception can be managed and permanent harm almost always avoided. If an unacceptable level of risk after vaccination is identified, administration of the vaccine can be halted almost immediately and the risk to future populations eliminated, providing for high institutional ability to mitigate: the committee estimated that over 99% of the time RRV-TV-induced adverse effects (once detected) could be mitigated.

USING THE RISK CHARACTERIZATION TO SUPPORT DECISION-MAKING

Table 3-1 highlights the differences in public-health consequences that could be expected if the vaccine is withdrawn or left on the market. If the vaccine remained on the market, the number of children who would suffer adverse health effects—ranging from diarrhea to death—from rotavirus *and* RRV-TV would be substantially less than the number who would suffer similar effects of rotavirus disease if the vaccine were removed from the market. If the vaccine remained on the market, most of the adverse effects would result from rotavirus disease in the unvaccinated population, *not* from the vaccine. If the vaccine were assumed to be more effective and used more widely, the number of deaths and illnesses from rotavirus would decrease, and those from vaccine-induced effects would increase. However, examining the ranges of effects in Table 3-1 indicates that such changes in assumptions would result in

even fewer adverse health effects with the vaccine on the market. The table also indicates that the vaccine-related risks are controllable, and adverse effects from the vaccine are readily detectable and treated. On the basis of this risk characterization alone, it appears that a case could be made to retain the vaccine on the market.

However, as described in Chapter 2, the risk characterization conducted here is designed to capture only the relatively direct public-health consequences associated with different decisions. This information should be relevant and important to FDA decision-making, but many other factors outside the scope of the risk characterization offered here are also relevant and should be considered. For example, the public has a high standard for and expectation of vaccine safety, and that public confidence in the safety of vaccines is important for public health and has implications extending far beyond the case of a single vaccine. A much more detailed analysis could be conducted to evaluate those indirect health consequences. Such an evaluation would require an even greater number of assumptions and estimates, such as whether (and the degree to which) confidence in vaccines would be eroded by the retention of RRV-TV, how erosion of confidence would affect the willingness of parents to vaccinate their children against other diseases, and whether and when those other diseases would increase in prevalence as a result of decreased immunizations.

A similar analysis could be conducted for situations in which the decision options being considered are more complex than the simple retain-withdraw options considered here. For mitigation-selection decisions that FDA would face, the options are likely to be much more nuanced than those described here. For example, mitigation options for any vaccine or drug could include changes in the recommended use or dosage or recommendations for preadministration screening or postadministration monitoring. Comparing more than two mitigation options by using the framework simply requires FDA to define each mitigation option clearly, to characterize the consequences of each by using the attributes, and to follow a logic similar to that illustrated in this case study. An attribute table similar to Table 3-1 would summarize the outcomes of all decisions being considered on a common basis (in terms of the same set of attributes). The decision-making step remains complicated by additional factors, as discussed above, but the use of this framework will enable FDA to compare the public-health consequences of multiple mitigation approaches on a common basis and to describe the consequences with a common language.

In the actual events on which this case study was based, the decision of whether the vaccine remained on the market was made by the manufacturer, not FDA. The manufacturer had its own set of decision criteria that may have included some of the risk-characterization information above but almost certainly included other considerations, including liability and financial issues. Research on the rates of intussusception and the risk associated with RRV-TV continued after 1999. In the interest of completeness, the committee reviewed data that are available now but were not when a decision about withdrawal of RRV-TV had to be made. In particular, later assessments of the risk of RRV-TV-induced intussusception yielded reduced risk estimates of one case in every 10,000-32,000 vaccinations (Murphy et al. 2003a; Murphy et al. 2003b). The mean annual rate of intussusception in the absence of vaccine was recently published in a retrospective analysis involving infants in three children's hospitals over a 5-year period. The mean annual intussusception rate was found to be 49.3 cases per 100,000 live births (inpatient cases, 27.1 cases per 100,000 live births; short-stay or emergency-department cases, 22.3 cases per 100,000 live births) with a case-fatality of about 1.3% (2 of 156 cases) (Cortese et al. 2009).

REFERENCES

AAP (American Academy of Pediatrics). 1998. Prevention of rotavirus disease: Guidelines for use of rotavirus vaccine. Pediatrics 102(6):1483-1491.

Alpert, J.J. 1999. PedComm: AAP Member Alert, July 15, 1999. American Academy of Pediatrics [online]. Available: http://www.westchestergov.com/Health/PCaaprecomm.htm [accessed Oct. 26, 2010].

Bernstein, D.I. 2009. Rotavirus overview. Pediatr. Infect. Dis. J. 28(Suppl. 3):S50-S53.

Braun, M.M. 2006. Vaccine Adverse Event Reporting System (VAERS): Usefulness and Limitations. Institute for Vaccine Safety, John Hopkins Bloomberg School of Public Health [online]. Available: http://www.vaccinesafety.edu/VAERS.htm [accessed Mar. 18, 2011].

CDC (U.S. Centers for Disease Control and Prevention). 1999a. Rotavirus vaccine for the prevention of rotavirus gastroenteritis among children: Recommendations of the Advisory Committee on Immunization Practices (ACIP). MMWR 48 (RR-2):1-23 [online]. Available: http://www.cdc.gov/mmwr/preview/mmwrhtml/00056669.htm [accessed Oct. 26, 2010].

CDC (U.S. Centers for Disease Control and Prevention). 1999b. Intussusception among recipients of rotavirus vaccine -United States, 1998-1999. MMWR 48(27):577-581 [online]. Available: http://www.cdc.gov/mmwr/preview/mmwrhtml/mm4827a1.htm [accessed Oct. 27, 2010].

CDC (U.S. Centers for Disease Control and Prevention). 1999c. Withdrawal of rotavirus vaccine recommendation. MMWR 48(43):1007 [online]. Available: http://www.cdc.gov/mmwr/preview/mmwrhtml/mm4843a5.htm [accessed Oct. 27, 2010].

CDC (U.S. Centers for Disease Control and Prevention). 2001. Immunization Coverage in the U.S.: January-December 1998 Table Data [online]. Available: http://wwwtest.cdc.gov/vaccines/stats-surv/nis/data/tables_1998.htm [accessed Oct. 27, 2010].

CDC (Centers for Disease Control and Prevention). 2009. Prevention of rotavirus gastroenteritis among infants and children recommendations of the Advisory Committee on Immunization Practices (ACIP). MMWR 58(RR02):1-25[online]. Available: http://www.cdc.gov/mmwr/preview/mmwrhtml/rr5802a1.htm [accessed Mar. 15, 2011].

Clemen, R.T. 1996. Making Hard Decisions: An Introduction to Decision Analysis. Belmont, CA: Duxbury Press. 664 pp.

Cortese, M.M., M.A. Staat, G.A. Weinberg, K. Edwards, M.A. Rice, P.G. Szilagyi, C.B. Hall, D.C. Payne, and U.D. Parashar. 2009. Underestimates of intussusception rates among U.S. infants based on inpatient discharge data: Implications for monitoring the safety of rotavirus vaccines. J. Infect. Dis. 2000(Suppl. 1):S264-S270.

Dennehy, P.H. 2008. Rotavirus vaccines: An overview. Clin. Microbiol. Rev. 21(1):198-208.

Ellenberg, S.S., and R.T. Chen. 1997. The complicated task of monitoring vaccine safety. Public Health Rep. 112(1):10–21.

Fischer, T.K., C. Viboud, U. Parashar, M. Malek, C. Steiner, R. Glass, and L. Simonsen. 2007. Hospitalizations and deaths from diarrhea and rotavirus among children <5 years of age in the United States, 1993-2003. J. Infect. Dis. 195(8):1117-1125.

Glass, R.I., P.E. Kilgore, R.C. Holman, S. Jin, J.C. Smith, P.A. Woods, M.J. Clarke, M.S. Ho, and J.R. Gentsch. 1996. The epidemiology of rotavirus diarrhea in the United States: Surveillance and estimates of disease burden. J. Infect. Dis. 174(Suppl. 1):S5-S11.

Joensuu, J., E. Koskenniemi, X.L. Pang, and T. Vesikari. 1997. Randomised placebo-controlled trial of rhesus-human reassortant rotavirus vaccine for prevention of severe rotavirus gastroenteritis. Lancet 350(9086):1205-1209.

Joensuu, J., E. Koskenniemi, and T. Vesikari. 1998. Symptoms associated with rhesus-human reassortant rotavirus vaccine in infants. Pediatr. Infect. Dis. J. 17(4):334-340.

Keefer, D.L., and S.E. Bodily. 1983. Three-point approximations for continuous random variables. Manage. Sci. 29(5):595-609.

Kramarz, P., E.K. France, F. Destefano, S.B. Black, H. Shinefield, J.I. Ward, E.J. Chang, R.T. Chen, D. Shatin, J. Hill, T. Lieu, and J.M. Ogren. 2001. Population-based study of rotavirus vaccination and intussusception. Pediatr. Infect. Dis. J. 20(4):410-416.

Lieu, T.A., M. Kulldorff, R.L. Davis, E.M. Lewis, E. Weintraub, K. Yih, R. Yin, J.S. Brown, and R. Platt. 2007. Real-time vaccine safety surveillance for the early detection of adverse events. Med. Care 45(10 Suppl. 2):S89-S95.

Murphy, T.V., P.M. Gargiullo, M.S. Massoudi, D.B. Nelson, A.O. Jumaan, C.A. Okoro, L.R. Zanardi, S. Setia, E. Fair, C.W. LeBaron, B. Schwartz, M. Wharton, and J.R. Livengood. 2001. Intussusception among infants given an oral rotavirus vaccine. N. Engl. J. Med. 344(8):564-572.

Murphy, B.R., D.M. Morens, L. Simonsen, R.M. Chanock, J.R. La Montagne, and A.Z. Kapikian. 2003a. Reappraisal of the association of intussusception with the licensed live rotavirus vaccine challenged initial conclusions. J. Infect. Dis. 187(8):1301-1308.

Murphy, T.V., P.J. Smith, P.M. Gargiullo, and B. Schwartz. 2003b. The first rotavirus vaccine and intussusception: Epidemiological studies and policy decisions. J. Infect. Dis. 187(8):1309-1313.

Parashar, U.D., R.C. Holman, K.C. Cummings, N.W. Staggs, A.T. Curns, C.M. Zimmerman, S.F. Kaufman, J.E. Lewis, D.J. Vugia, K.E. Powell, and R.I. Glass. 2000. Trends in intussusception-associated hospitalizations and deaths among U.S. infants. Pediatrics. 10696:1413-1421.

Perez-Schael, I., M.J. Guntinas, M. Perez, V. Pagone, A.M. Rojas, R. Gonzalez, W. Cunto, Y. Hoshino, and A.Z. Kapikian. 1997. Efficacy of the rhesus rotavirus-based quadrivalent vaccine in infants and young children in Venezuela. N. Engl. J. Med. 337(17):1181-1187.

Rennels, M.B., R.I. Glass, P.H. Dennehy, D.I. Bernstein, M.E. Pichichero, E.T. Zito, M.E. Mack, B.L. Davidson, and A.Z. Kapikian. 1996. Safety and efficacy of high-dose rhesus-human reassortant rotavirus vaccines—report of the National Multicenter Trial. Pediatrics 97(1):7-13.

Rennels, M.B., U.D. Parashar, R.C. Holman, C.T. Le, H.G. Chang, and R.I. Glass. 1998. Lack of an apparent association between intussusception and wild or vaccine rotavirus infection. Pediatr. Infect. Dis. J. 17(10):924-925.

Santosham, M., L.H. Moulton, R. Reid, J. Croll, R. Weatherholt, R. Ward, J. Forro, E. Zito, M. Mack, G. Brenneman, and B.L. Davidson. 1997. Efficacy and safety of high-dose rhesus-human reassortant rotavirus vaccine in Native American populations. J. Pediatr. 131(4):632-638.

Tucker, A.W., A.C. Haddix, J.S. Bresee, R.C. Holman, U.D. Parashar, and R.I. Glass. 1998. Cost-effectiveness analysis of a rotavirus immunization program for the United States. JAMA. 279(17):1371-1376.

4

Case Study of a Targeting Decision

This chapter describes a case study that uses the risk-characterization framework to evaluate several food categories and the potential human-health consequences of foodborne illnesses associated with them. The case study was selected because it is an example of the kind of evaluation and analysis that might be used to support a targeting decision, it is relevant to a scenario provided to the committee by the Food and Drug Administration (FDA), and several committee members had relevant expertise. As noted in Chapter 2, targeting decisions are the most closely related to risk-ranking questions that are of interest to some FDA centers. The data used were gleaned from publicly available Web sites or publications or provided by FDA. The committee did not conduct exhaustive literature searches or reviews, and all information is illustrative. The case study simply illustrates how the committee's framework might be used for a targeting decision.

FRAMING THE ISSUE: FOOD SAFETY

FDA has the responsibility for ensuring the safety of about 80% of the U.S. food supply and regulates most foods and food ingredients, which range from raw commodities to highly processed foods (GAO 2008). The extent of FDA oversight is variable and ranges from relatively little public oversight of some products, such as fresh vegetables, to a highly regulated system for others, such as canned foods. The diversity and number of food products and varied regulations pose challenges to any attempt to characterize or rank food categories on the basis of health consequences. The task is complicated by the rapid adjustments of the global food system, which changes with market access and availability. Several recent reports have identified important gaps and deficiencies in the oversight of foods by the federal government (GAO 2010; IOM/NRC 2010), and a report recently released by the Institute of Medicine (IOM/NRC 2010) described and compared a number of risk-ranking models for foods (see Box 4-1).

DECISION CONTEXT FOR THE CASE STUDY

The decision context considered in this case study was one of allocating food-safety inspection resources; that is, if additional inspection resources were available, how should they be allocated among the various food categories to maximize public-health protection? That general decision problem is too large to be undertaken as a case study, so a much simplified decision context and evaluation were selected. Rather than considering all the different options for what types of food could be inspected, this case study considers only three specific food categories: leafy greens, shrimp, and canned foods. The categories were chosen to highlight products that are inherently different with respect to level of

BOX 4-1 Risk-ranking Models for Foods

Among FDA-regulated products, foods are probably the category that has been the focus of more risk-assessment and ranking studies than any other products. Several government agencies and research groups have developed risk-ranking approaches and models, and the National Research Council (NRC) and the Institute of Medicine (IOM) have published several reports addressing the need for ensuring the safety of foods (IOM/NRC 1998, 2003, 2010). In the latest report, different risk-ranking models developed by researchers and regulators in different countries were contrasted. The "degree of complexity, level of quantification, and approach to model construction" differed among the models (IOM/NRC 2010). However, some of the common criteria identified were "(1) burden of illness…(2) illness severity, (3) population susceptibility, (4) likelihood of contamination, (5) potential for agent amplification, and (6) breadth of exposure" (IOM/NRC 2010, p. 87). The risk-characterization framework proposed here considers all those factors although some are not called out explicitly: (3) and (6) are included in exposed population; (1) and (2) are covered and expanded under mortality and adverse health effects; (4) and (5) must be considered to develop estimates of the number of deaths and other adverse health effects and thus are implicitly included in this framework. However, like the risk-ranking models described in the recent IOM/NRC report, the risk-characterization framework proposed here was designed for a specific purpose, which was to characterize the public-health consequences of various decisions at FDA among all its programs, and therefore does not include program-specific attributes that would not be universally applicable, such as the probability of contamination and amplification in the food supply.

processing, origin, and potential risks. Furthermore, rather than identifying all the possible allocations of inspection resources that could potentially be compared, this case focuses only on characterizing the public-health consequences associated with each food category assuming the current regulatory and inspection regime. The results of this evaluation could be used directly for ranking or comparing the food categories on the basis of risk or could serve as input into decisions for allocating resources among the three food categories to maximize protection of the public from foodborne illnesses. Possible extensions of the analysis that would support resource-allocation decisions more directly are described in the final section of this chapter.

CHARACTERIZING THE PUBLIC-HEALTH CONSEQUENCES

Data on foodborne illnesses are generally available from the Centers for Disease Control and Prevention (CDC), but estimating the number of illnesses caused by the food categories evaluated in this case study required additional data and various assumptions. The committee consulted three primary sources of data—CDC (2010), Scallan et al. (2011a,b), and Hoffmann et al. (2007)—to estimate the number and severity of foodborne illnesses associated with each food category. CDC (2010) provides data on reported outbreaks of foodborne illnesses caused by various food commodities by pathogen for 1998-2007. Data are also reported on foodborne illnesses of unknown etiology. The actual cases of foodborne illness, however, are thought to exceed the number of reported cases associated with outbreaks substantially (Mead et al. 1999; Scallan et al. 2011a), so the reported cases represent only a fraction of the actual cases. Scallan et al. (2011a) provided pathogen-specific estimates of under-reporting and under-diagnosis that range from a factor of 2 to factors over 700.[1] Scallan et al. (2011a,b) also provided

[1]Under-reporting of illnesses that tend to be severe, such as those associated with botulism poisoning, is minimal whereas under-reporting of illnesses that tend to be less severe, such as those associated with *Salmonella*, is much greater.

estimates of the total number of cases of foodborne illness, hospitalizations, and deaths in the United States annually attributed to specific pathogens and those from unspecified agents. However, the data provided by Scallan et al. do not attribute the cases to specific food groups or food types. Hoffmann et al. (2007) described an expert elicitation study in which 42 nationally recognized food-safety experts were asked to provide estimates of the percentage of cases of foodborne illness caused by a specific pathogen that are attributable to 11 food categories. The committee combined the number of cases by pathogen from Scallan et al. (2011a) with the estimated percentage of illnesses from that pathogen attributed to specific food groups from Hoffmann et al. (2007) to estimate the annual number of cases from consumption of each food category. In each of the cases, some additional assumptions were necessary to extrapolate from the food categories of Hoffmann et al. to the more specific, narrower categories evaluated in this case study. CDC outbreak data were examined to support the estimates developed by the committee.

The following discussion provides some background information on the three food categories selected for evaluation and describes the basis of the risk characterization. The attributes and their estimates are summarized in Table 4-1.

Domestic Leafy Greens

Domestic leafy greens are widely consumed in the United States; although many of the leafy greens are available in frozen or canned form, the fresh product dominates the market. For example, fresh spinach made up 65.5% of total per capita spinach consumption in 2008, an increase from 58.5% in 1990 and from only 17.0% in 1970 (USDA 2010). Today, annual per capita consumption of fresh leafy greens is about 29 lb; romaine lettuce, leaf lettuce, head lettuce, and spinach make up most of the fresh product. However, other, less well-known greens—such as arugula, radicchio, and mizuna—have increased in consumption, often as components in fresh bagged mixed salads.

Most of the fresh greens are grown domestically. In 2006, only 3% of spinach, 2% of head lettuce, and 1% of leaf and romaine lettuce were imported for domestic consumption (Calvin et al. 2009). Monterey, Santa Clara, and San Benito Counties produce more than half the U.S. fresh market for spinach and supply up to 80% of other leafy greens sold in the United States (Calvin et al. 2009). Products are widely and quickly shipped around the country. Although the greens are often shipped directly from the field, they may be mixed and repackaged by processors before being shipped to retail and food-service outlets.

Fresh leafy greens are often consumed raw with little preparation, and the shift toward fresh products has increased the associated food-safety risk because products consumed raw have not been treated with heat or other kill steps. Outbreaks associated with leafy greens were 38.6% more frequent in 1996-2005 than in 1986-1995 (Mandrell 2009). However, there does not seem to be a well-identified cause of outbreaks associated with leafy greens: risk factors are numerous and include the potential for contamination before and after harvest (Mandrell 2009).

Although reported outbreaks associated with leafy greens are relatively rare, there have been some major ones. In 2006, for example, an outbreak of *Escherichia coli* O157:H7 in spinach resulted in illness in consumers in 26 states; of 204 cases, 31 cases of hemolytic-uremic syndrome (a serious complication of *E. coli* O157:H7 exposure) and three deaths were reported and attributed to the exposure (Calvin 2007). Mandrell (2009) noted a review by CDC that found that leafy greens were associated with 502 outbreaks, more than 18,000 cases of illness, and 15 deaths from 1973 to 2006.

TABLE 4-1 Risk Attributes for Targeting Decision for Foods

Attribute	Metric	Food Category		
		Domestic Leafy Greens	Shrimp	Canned food
Exposed population	Number exposed annually (number of U.S. population that consume product)	294 million (232-304 million)	263 million (155-279 million)	308 million (307-310 million)
	Populations of concern	Children under 5 years old, the elderly, and the immunosuppressed are more susceptible and suffer more severe effects from foodborne illnesses; pregnant women are of special concern for specific pathogens, such as *Listeria*.		
Mortality	Number of deaths per year	280 (5-590) total: 60 (5-150) from known pathogens 220 (0-440) from unspecified agents	56 (0-110) total: 10 (0-19) from known pathogens 46 (0-91) from unspecified agents	0.2 (0.1-2.8)
Morbidity	Number experiencing severe adverse health effects per year	13,900 (280-28,200) total: 4,600 (280-7,900) from known pathogens 9,300 (0-20,300) from unspecified agents	2,730 (15-6,700) total: 730 (15-1,200) from known pathogens 2,000 (0-5,500) from unspecified agents	20 (20-400)
	Number experiencing less severe adverse health effects per year	165,000 (1,500-266,000) total: 35,000 (1,500-56,000) from known pathogens 130,000 (0-210,000) from unspecified agents	34,800 (120-54,500) total: 6,800 (120-10,500) from know pathogens 28,000 (0-44,000) from unspecified agents	110 (110-200)
	Number per year experiencing adverse health effects that affect only quality of life	6.3 million (58,000 to 9.9 million) total: 1.3 million (58,000 to 2 million) from known pathogens 5 million (0 to 7.9 million) from unspecified agents	1.6 million (7,000 to 2.4 million) total: 255,000 (7,000-390,000) from known pathogens 1.3 million (0 to 2 million) from unspecified agents	4,100 (4,100-84,000)
Personal controllability	Degree to which a person can eliminate or reduce his or her own risks through voluntary actions	40-50% of cases of foodborne illnesses from leafy greens could be eliminated or reduced by personal action by the consumer	10-15% of cases of foodborne illnesses from shrimp could be eliminated or reduced by personal action by the consumer	45-60% of cases of foodborne illnesses from commercially canned foods could be eliminated or reduced by personal action by the consumer

TABLE 4-1 Continued

Attribute	Metric	Food Category		
		Domestic Leafy Greens	Shrimp	Canned food
Ability to detect adverse health effects	Ability of informed institution to detect population-level effects associated with product being evaluated	3% or fewer of all cases of foodborne illness caused by leafy greens could be detected and successfully attributed.	3-5% of all cases of foodborne illness caused by shrimp could be detected and successfully attributed.	About 50% of cases of botulism from commercial canned foods could be detected; 3% or fewer of all other types of foodborne illness caused by commercially canned foods could be detected and successfully attributed.
Ability to mitigate adverse health effects	Probability that an informed institution will be able to reduce or mitigate adverse health effects associated with the product being evaluated if such a problem is known to exist	<10%	10-50%	50-75%

Exposed Population

Although most people 2 years old and older consume lettuce or fresh leafy greens at some point in a year, the average daily consumption of greens is relatively low, less than 0.1 cup per day (NCI 2010). On the basis of nationally representative data, Tooze et al. (2006) estimated that 48% of men and 57% of women consume dark-green vegetables in a day. Therefore, although the average amount consumed per day may be small, the population exposed over a year is potentially large. For purposes of this case study, the committee estimated that the number of people who consume leafy greens in a year could range from 75% to 98% of the U.S. population. The committee's best estimate is that 95% of the U.S. population consumes some leafy greens during the year and thus would be exposed to the potential for foodborne pathogens from this food product. Multiplying those percentages by an assumed U.S. population of 310 million (U.S. Census Bureau 2010) yields the estimates shown in Table 4-1.

Outbreaks associated with leafy greens have typically been attributed to microbiologic contamination (Beuchat 1996). Although some investigations have shown that the incidence of pathogens on greens is relatively low (Mandrell 2009), the infectious dose of some pathogens is also low, so the effect of even low pathogen concentrations on leafy greens can be important. In general, infants and young children, older people, and immunosuppressed populations are more susceptible to foodborne illnesses and are likely to suffer more severe effects if they contract the illnesses. Accordingly, those groups are highlighted as populations of special concern.

Mortality and Morbidity

To estimate the mortality and morbidity attributable to leafy greens, the committee used a variety of data to make two types of estimates:

- The annual number of cases of foodborne illness caused by leafy greens attributable to specific pathogens or of an unknown etiology. Those estimates were based on data from CDC (2010), Scallan et al. (2011a,b), and Hoffmann et al. (2007).
- The rates of death and hospitalization for foodborne illness caused by each pathogen and of unknown etiology. Those rates were derived from data in Scallan et al. (2011a,b).

The numbers of cases and the rates of death and hospitalizations were used to calculate the estimated numbers of deaths and other adverse health effects summarized in Table 4-1. Figure 4-1 illustrates the general approach taken by the committee.

Annual Number of Cases of Foodborne Illnesses Associated with Leafy Greens

As noted above, the committee used three primary sources to estimate the annual number of cases of foodborne illness caused by leafy greens. An estimate of the average annual number of reported cases attributed to various pathogens (and of unknown etiology) and caused by consumption of leafy greens was developed directly from the CDC (2010) data. A second estimate of the number of cases of illness from leafy greens was developed by using data from Scallan et al. (2011a,b) and Hoffmann et al. (2007) as follows. Combining the total number of cases of foodborne illnesses by pathogen from Scallan et al. (2011a,b) with the attribution to food categories from Hoffmann et al. (2007) yielded an estimate of the annual number of cases from consumption of produce by pathogen. The CDC (2010) data were used to estimate the percentage of cases of illness caused by consumption of produce that were attributable to leafy greens by pathogen.

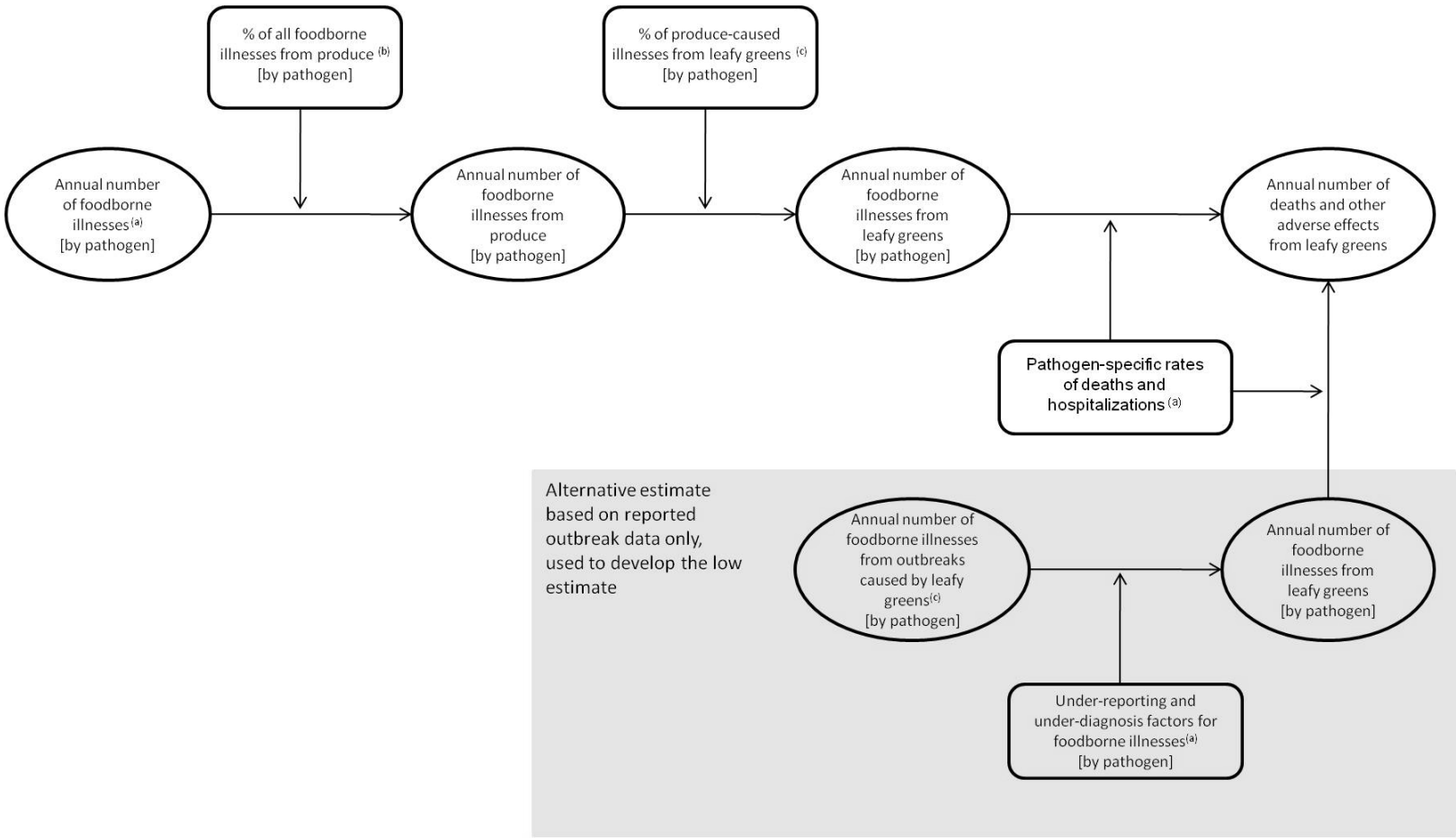

FIGURE 4-1 General approach used by the committee to estimate annual number of deaths and other adverse effects associated with leafy greens. (a) Scallan (2011a,b), (b) Hoffmann et al. (2007) , and (c) CDC (2010).

Rates of Foodborne Illnesses from Leafy Greens

From the data in Scallan et al. (2011a,b), hospitalization and death rates were calculated for each pathogen and for foodborne illnesses from unspecified agents. The rates and the "multiplier" for each pathogen that indicates the degree to which illnesses associated with a given pathogen are under-diagnosed and under-reported were used to estimate the number of deaths and adverse health effects of different severity from leafy greens, as described below.

Estimates of the Number of Deaths from and Other Adverse Health Effects of Leafy Greens

There is substantial uncertainty in the estimated total number of foodborne illnesses that occur annually in the United States and in how many of those illnesses are caused by leafy greens and by what pathogens. That uncertainty is reflected in the committee's approach described below to derive the low, high, and best estimates for the number of foodborne illnesses (by pathogen and from unspecified causes) from consumption of leafy greens.

- *Low Estimate*. The CDC (2010) data were used to calculate the reported illnesses for leafy greens. Those estimates were then scaled up with the pathogen-specific multipliers in Scallan et al. (2011a) to yield the pathogen-specific numbers of cases of illness from leafy greens. For the low estimate, no foodborne illnesses from unspecified agents were attributed to leafy greens.
- *Best Estimate*. The cases of illness attributable to produce were first calculated by using either the mean or modal estimates of the number of illnesses by pathogen as reported by Scallan et al. (2011a) and either the pathogen-specific or average expert-based attribution calculated from Hoffmann et al. (2007). Thirty percent of all the cases of foodborne illness from unspecified agents reported in Scallan et al. (2011b) were also included; this percentage is based on the overall percentage of total foodborne illnesses attributed to produce in Hoffmann et al. (2007). Next, the cases attributed to leafy greens were calculated by using either the pathogen-specific or the average attributions from all produce-caused illnesses that could be calculated from the CDC (2010) data.
- *High Estimate*. Scallan et al. (2011a,b) provided not only the mean or modal estimates of the number of foodborne illnesses but a 90% credible interval for those numbers. The high estimate was calculated analogously to the best estimate except that the high end of the range provided by Scallan et al. was used rather than the mean or modal estimates.

For all three estimates (low, best, and high estimates), the pathogen-specific hospitalization and death rates derived from Scallan et al. (2011a,b) were used to calculate numbers of severe health effects (assumed to be hospitalizations) and deaths. Generic rates derived from Scallan et al. were used for illnesses for which no pathogen-specific rates were available. Illnesses that did not lead to death or to hospitalization were classified as producing less severe adverse effects or effects that were related only to quality of life. Those illnesses were apportioned to the two morbidity categories by using the under-diagnosis and under-reporting factors from Scallan et al. (2011a,b). For example, if *Salmonella* is assumed to have a combined factor of 29 for under-diagnosis and under-reporting, the committee assumed that there will be about 29 times as many illnesses that affect only quality of life compared with those which will give rise to "less severe adverse effects."

The committee notes that for the best and high estimates the vast majority of the cases of foodborne illness are from unspecified agents. Of the pathogen-specific illnesses, a large proportion are attributed to *Norovirus*, as described in Mead et al. (1999) and updated in Scallan et al. (2011a); such illnesses may not typically be thought of as foodborne.

Personal Controllability

People consuming leafy greens at home have several options available to reduce their risks of foodborne illnesses from those greens: primarily by refrigerating foods at proper temperatures; checking for signs of spoilage; thoroughly washing hands, utensils, and the greens before use; and maintaining strict separation of the greens from utensils used to prepare other foods, especially raw meats (FSIS 2006). People consuming leafy greens in restaurants, in contrast, have little ability to reduce their risks of contracting a foodborne illness other than not ordering or consuming foods that contain raw leafy greens. About 27% of all meals consumed in the United States are prepared outside the home (Lin et al. 1998). Lacking any data to the contrary, the committee assumed that leafy greens were just as likely to be consumed at home as away from home, so at least 27% of leafy greens consumed are prepared outside the home, and at least 27% of the cases of foodborne illnesses from leafy greens are not personally controllable by the individual consumer.

The cases remaining arise from home use. Although those cases are theoretically controllable, people must have knowledge that they are potentially at risk and knowledge of the steps that they can take to reduce those risks for them to have practical controllability over their risks. Reports on consumers' behavior before, during, and after the 2006 spinach outbreak show confusion about details of the recall announcement (for example, which spinach products were involved) and of the associated symptoms of *E. coli* O157:H7-related illness (Hallman et al. 2009); this suggests that consumers have imperfect knowledge of how to prevent illnesses from leafy greens. Even during the recall with considerable media attention, only two-thirds of those surveyed reported that bloody diarrhea was a symptom of the infection. Furthermore, there may be some contamination that cannot be removed even if all proper precautions and procedures are followed. Given the uncertainty surrounding consumer knowledge and the ability to remove contamination, the committee estimated that 55-70% of those preparing leafy greens at home have practical control over the risks. Given the estimate that 73% of cases of foodborne illnesses from leafy greens result from home preparation, about 40-50% of the total cases of illnesses from leafy greens are personally controllable.

Ability to Detect Adverse Health Effects

It is difficult in general to detect foodborne illness and to attribute it correctly to a specific cause, especially foodborne illness that occurs sporadically as opposed to illness that is associated with a specific outbreak. Use of surveillance data and advanced genotyping techniques has improved identification of outbreaks and the ability to identify and trace back potentially contaminated products associated with them. However, sporadic cases of foodborne illness occur much more frequently and in far greater numbers than cases associated with outbreaks (Mead et al. 1999).

The types of foodborne illnesses associated with leafy greens—such as those caused by *Norovirus*, *Campylobacter*, and *Salmonella*—are also those estimated to be largely under-reported (Mead et al. 1999); incidents of illnesses are far more likely to occur as individual cases than as concentrated outbreaks, and this compounds the difficulty in detecting those cases. On the basis of the estimated under-reporting, the committee estimates that fewer than 3% of all cases of foodborne illnesses from leafy greens can be detected and successfully attributed to the causal agent.

Ability to Mitigate Adverse Health Effects

In food safety, risk mitigation often refers to prevention or intervention—that is, preventing contamination or treating to ensure that contamination is reduced. In 1998, FDA published voluntary guidelines (Good Agricultural Practices) to help growers reduce the risk of contamination on the farm. Since the 2006 outbreak linked to spinach, the California leafy-greens industry, through the Leafy Greens

Marketing Agreement (LGMA), developed best-practice standards to guide production practices and control and monitoring of produce. The preharvest and harvest handling have been identified as key points where hazards may be introduced through exposure to environmental and production hazards, such as contaminated water, manure, and poor field sanitation. However, postharvest handling, produce mixing in fresh packaging, and distribution are also important. Although FDA and others have developed guidance to reduce health risks, control of product movement and the ability to prevent contamination rest with growers, processors, retailers, food-service industries, and ultimately consumers (as discussed above in connection with personal controllability). As noted, testing of produce has increased throughout the production process (at time of field harvest, during initial processing, and by buyers and distributors) in the private sector, and the effectiveness of the activities is reflected in the overall rates of foodborne illnesses from leafy greens.

In this framework, the ability to mitigate refers to the ability to manage, reduce, or otherwise control any adverse health effects of the products being evaluated, assuming that such effects occur and are detected. Because of the complexity of the food supply and the diversity of risks associated with foods, mitigation of adverse health effects from foodborne illness is challenging. At an institutional level, mitigation efforts could include efforts to prevent additional cases from occurring through recalls of tainted products and efforts to reduce disease symptoms through outreach and treatment education. For leafy greens (and other food products), prevention of additional cases rarely occurs because of the inability to identify the food vehicle. For a highly perishable food product with a short shelf-life, such as leafy greens, even when a problem has been detected most of the product has already been consumed, and this limits further the ability of institutions to mitigate adverse health effects. On the basis of those factors, the committee estimated that the probability that institutions can mitigate adverse effects from foodborne illness caused by consumption of leafy greens through effective recalls or better treatment after a potential problem has been detected is less than 10%.

Shrimp

About 90% of the shrimp consumed in the United States is imported (NOAA 2010). According to the U.S. National Marine Fisheries Service, the United States imported more than 548,000 metric tons of shrimp in 2009 (NOAA 2010). The increase in shrimp importation mirrors a dramatic increase in overall seafood importation. The United States receives 318,000 tons of its imported shrimp from Thailand, Ecuador, and China (NOAA 2010). Other countries that supply shrimp to the United States include Indonesia, Bangladesh, Mexico, Vietnam, and India (FAO 2010). Most of the shrimp imported into the United States are fresh or frozen whole. Recently, there has been a dramatic increase in imported breaded shrimp; however, this product accounted for less than 10% (37,427 metric tons) of shrimp imports in 2009 (NOAA 2010).

Frozen shrimp imported into the United States are directly transported, further processed, or placed into cold storage. Directly transported shrimp products are delivered to restaurants or grocery chains. Further-processed shrimp products are delivered to other restaurants or local grocery chains. Shrimp products can remain in cold storage for long periods. It is said that 80% of imported shrimp is consumed in restaurants and the rest sold in grocery stores (R. Fischer, FDA, personal communication, March 2010).

In 2005-2009, nearly 100,000 lines[2] of shrimp were imported into the United States (mostly raw and frozen); 2,030 lines were refused. Major reasons for refusal included evidence of filth (21%), *Salmonella* (16%), and drug residues (23%). Those refusal data are similar to published data on the 2001-2003 experience (Wan Norhana et al. 2009). Most of the reported outbreaks associated with shrimp for

[2]A *line* is defined by FDA as "each portion of an entry which is listed as a separate item on an entry document. An importer may identify merchandise in an entry in as many portions as he chooses, except each item in the entry having a different tariff description and rate must be listed separately" (FDA 2010a).

which an infectious agent has been identified have been attributed to *Salmonella* contamination, and salmonellosis has been linked to shrimp from aquaculture ponds (Koonse et al. 2005). Consequently, the United States has a zero-tolerance policy regarding *Salmonella* (that is, it requires the absence of *Salmonella*) for raw or cooked and ready-to-eat shrimp (Koonse et al. 2005; Wan Norhana et al. 2009).

Exposed Population

The per capita U.S. consumption of shrimp increased from 2.4 kg in 1997 to a peak of 4.6 kg in 2006 (The Fish Site 2008). It is estimated that 85% of the U.S. population consumes some shrimp each year (IOM 2007). For purposes of this case study, it is assumed that 50-90% of the U.S. population consumes shrimp; the best estimate is 85%. Combining those estimates with a U.S. population size of 310 million yields the estimates of the exposed population shown in Table 4-1.

As described above in the discussion of leafy greens, infants and young children, older people, and immunosuppressed populations are more susceptible to foodborne illnesses generally and are more likely to suffer more severe effects if they do contract the illnesses. Accordingly, those groups are highlighted as populations of special concern.

Mortality and Morbidity

According to outbreak reports to CDC, 62 outbreaks of illness were linked to shrimp in 2001-2005, and they were associated with 618 cases of illness (see Table 4-2). Cases of illness in 2001-2005 ranged from 25 to 184 per year (Tsutumi 2007). Shrimp outbreaks can be confounded because shrimp are often mixed with various other food ingredients, and this complicates investigations that are trying to identify outbreak sources. Most outbreaks cannot be linked to an identified pathogen or agent, as shown in Table 4-2; the data in the table reflect only reported outbreaks and cases.

TABLE 4-2 Agents Linked to Shrimp-Associated Outbreaks and Cases in the United States, 2001-2005

Infectious Agents	Number of Outbreaks	Number of Cases of Illness
Clostridium perfringens	0	0
Shigella sonnei	1	2
Salmonella spp.	5	58
Staphylococcus spp.	0	0
Vibrio parahaemolyticus	4	121
Multiple bacteria	0	0
Norovirus/Norwalk viruses	3	182
Unknown	49	255
Total	*62*	*618*

Source: Adapted from Tsutumi 2007.

Mortality and morbidity attributable to shrimp were calculated by using the methods described above for leafy greens and data from Tsutumi (2007), CDC (2010), Scallan et al. (2011a,b), and Hoffmann et al. (2007). CDC (2010) reported data on illnesses associated with shrimp and seafood, and Hoffmann et al. (2007) provided estimates of the percentages of cases of foodborne illness from specific pathogens attributable to seafood (that is, seafood was one of the food categories evaluated in the expert-elicitation study). Specifically, the values shown in Table 4-1 were estimated as follows:

- *Low Estimate*. The low estimate was calculated by using the outbreak data shown in Table 4-2. Those estimates were then scaled up by using the pathogen-specific multipliers in Scallan et al. (2011a) to yield the pathogen-specific numbers of cases of illness. Numbers of deaths, severe health effects, less severe health effects, and adverse quality-of-life health effects were then calculated according to the general method used for leafy greens.
- *Best Estimate*. Cases of illness attributable to seafood were first calculated by using the data on pathogen-specific illnesses reported by Scallan et al. (2011a) and the pathogen-specific or average expert-based attribution from Hoffmann et al. (2007). Twenty-five percent of all the cases of foodborne illnesses from unspecified agents as reported in Scallan et al. (2011b) were also included; this percentage is based on the overall percentage of cases attributed to seafood in Hoffmann et al. (2007). Next, the cases attributed to shrimp were calculated by using the pathogen-specific or average attributions that could be calculated from the CDC (2010) data. The numbers of death, severe health effects, less severe health effects, and adverse quality-of-life health effects were calculated according to the general method described for leafy greens.
- *High Estimate*. The high estimate was calculated analogously to the best estimate except that the high end of the range of number of foodborne illnesses provided in Scallan et al. (2011a,b) was used.

Like the estimates of the number of adverse health effects attributable to leafy greens, the best and high estimates for illness associated with shrimp include a large number of cases due to *Norovirus*.

Personal Controllability

The probability of contracting foodborne illnesses from shrimp can be largely minimized through proper hygiene in food preparation and cooking shrimp to an appropriate temperature; both activities can be practiced during home preparation of shrimp. However, as mentioned above, it is estimated that about 80% of shrimp consumed in the United States is consumed in restaurants rather than prepared at home. For food prepared in restaurants, the only action available to a person to minimize (or control) risks of foodborne illnesses from shrimp is not to order or consume it. Although that option is available, there is no reason, in the absence of any specific information about heightened risk from shrimp, for an individual consumer to believe that another food choice would carry lower risks and thus no reason to choose a different food as a way to control personal risks.

At most 20% of the total number of cases of foodborne illnesses from shrimp are potentially controllable by individual action on the basis of the percentage of shrimp consumed at home. Following the same logic described for leafy greens, the committee assumes that 60-80% of those consuming shrimp at home have sufficient knowledge to be able to reduce the likelihood of foodborne illnesses, resulting in an estimate that about 10-15% of all cases of foodborne illnesses from shrimp are personally controllable, that is, could be avoided by the actions of individual consumers.

Ability to Detect Adverse Health Effects

Because a variety of pathogens and agents can be linked to contaminated shrimp, the disease manifestations are varied. Hence, an institution's ability to detect a problem is based on common symptoms and requires a clinic or hospital visit where appropriate diagnostic tests are performed. On the basis of studies that examined a health department's ability to detect a contamination problem, at least five or six patients would need to be identified to establish a possible link to a contaminated food item (Bender et al. 1997; Rounds et al. 2010). With appropriate diagnostics and reporting to a public-health official, baseline epidemiologic features can be determined and can generate hypotheses as to likely sources of illness. Mild cases or cases with long incubation periods are not likely to be reported or linked to a specific food item, nor are sporadic cases, which make up the vast majority of cases of foodborne illness.

As in the case of leafy greens, most foodborne illnesses associated with shrimp have high estimated rates of under-reporting. On the other hand, because most shrimp is consumed in restaurants, there may be a higher likelihood that illnesses will be reported and a higher likelihood that multiple people will be affected at one time and an "outbreak" identified than when foods are prepared at home. On the basis of the under-reporting and under-diagnosis estimates in Scallan et al. (2011a) and the specific pathogens associated with shrimp, the committee estimated that about 3-5% of foodborne illnesses caused by shrimp would be detected.

Ability to Mitigate Adverse Health Effects

In the United States, the Federal Food, Drug, and Cosmetic Act gives FDA the responsibility of ensuring that no adulterated or misbranded food, including seafood, enters interstate commerce [21 USC §331]. The Office of Food Safety in FDA's Center for Food Safety and Applied Nutrition (CFSAN) oversees seafood safety. Its regulatory activities include control of foodborne pathogens and contaminants (such as methylmercury, *Vibrio*, and drug and chemical residues), inspection and compliance, and importation and exportation. Three other federal agencies work with FDA to protect the food supply system: the National Oceanic and Atmospheric Administration of the Department of Commerce conducts fee-for-service inspection for the industry and controls domestic fishing activities by prohibiting harvesting when the water is unsafe because of pollution or contamination, the Environmental Protection Agency sets tolerances for pesticide-residue limits and ensures cleanliness of air and water, and the U.S. Department of Agriculture is involved in seafood regulations and is responsible for promoting the aquaculture industry.

Under the current Hazard Analysis and Critical Control Points (HACCP) regulations, domestic and foreign seafood processors must have and implement a written HACCP plan and follow appropriate corrective action, verification, record-keeping, and training. The industry's implementation of the seafood HACCP regulation is investigated as a part of FDA's compliance programs. CFSAN provides the guidance for its field officers who inspect the industries. Domestic and foreign seafood processors are inspected for their compliance with HACCP regulations and other, non-HACCP attributes, such as filth and decomposition. The effectiveness of those programs in preventing contaminated products from reaching consumers is reflected in the estimates of the number of cases of foodborne illnesses that occur, as summarized in Table 4-1

The ability to mitigate a problem with shrimp if such a problem is detected and successfully attributed relies on the ability of appropriate institutions to remove the contaminated product from the marketplace through product recalls and consumer education. If the source of the problem is fresh shrimp, the situation is analogous to that described above for leafy greens: most of the product will have been consumed by the time the problem is identified. If the problem occurs in frozen and packaged shrimp, the ability to track the problem to specific lines and to recall the product successfully is somewhat greater because frozen shrimp may be stored for up to 2 years before being distributed to consumers. The

next section on canned foods includes a discussion of the effectiveness of recalls. Canned foods are characterized by a longer shelf-life than frozen shrimp, but it is unclear whether the longer shelf-life makes it easier or harder to recall a product from the market. For example, canned foods that remain on the market may be purchased months after a recall and stored by consumers for many more months or even years before being consumed. Frozen shrimp may be stored for relatively long periods before distribution; thus, if they are recalled, a higher percentage of the product may be in centralized storage than in consumers' homes at the time of the recall. Considering both fresh and frozen shrimp and the estimated ability to mitigate adverse health effects of consumption of leafy greens and canned foods described elsewhere in this case study, the committee estimated the ability of institutions to mitigate adverse health effects of shrimp consumption to be 10-50%.

Commercial Canned Foods

According to the Economic Research Service (ERS), the United States produces about 150 lb of domestic canned foods per capita per year (USDA 2009). They consist of a wide array of acidified and low-acid categories and include meats, vegetables, fruits, dairy foods, beverages, seafood, cereals, and multi-ingredient foods. Canned foods imported into this country are similar to domestic canned foods and also consist of a wide array of products. There is no available information on the geographic source distribution of nondomestic canned foods, but they are probably imported from at least 50 countries.

In the canning process, food products are subjected to a thermal treatment in sealed metal containers or cans to kill as many microorganisms present in the food in the form of vegetative cells or spores as possible. Because of the potential risk of botulism caused by improperly treated cans contaminated with spores of *Clostridium botulinum*, canned foods are among the most strictly regulated food categories. Modern canning technologies have been optimized thanks to years of research and strict regulations and typically are capable of delivering an equivalent bacterial killing of 12D or 12 log CFU of *C. botulinum* spores.

Recalls of canned foods are relatively unusual. Despite the large volume of product manufactured in the United States and overseas, there are rarely more than five recalls of canned foods in a year (FDA 2010b). From January to August 2010, there were three recalls; in 2009, only one recall was reported. The most common causes of recalls are under-processing, presence of allergens, and mislabeling, and they rarely involved a proven health risk.

According to FDA records, the shipments of imported canned foods have averaged about 97,240 lines in the last 5 years. About 2-3% of those lines were refused at the port of entry. The main reason for refusal was the lack of registration by the manufacturer; other frequent reasons were mislabeling and failure to file a process schedule with FDA.

Exposed Population

No data are available on the percentage of the U.S. population that regularly consumes canned foods. However, given the prevalence of these products in our food supply and the relatively high annual per capita production of 150 lb of canned foods (USDA 2009), the committee assumed that an overwhelming majority of Americans consume canned foods during the course of the year—that is, between 99-99.9% (best estimate, 99.5%), which accounts for people who prefer to consume nonprocessed food. Multiplying those percentages by an assumed U.S. population of 310 million yields the estimates shown in Table 4-1.

As described above in the discussion of leafy greens and shrimp, infants and young children, older people, and immunosuppressed populations are more susceptible to foodborne illnesses generally and are more likely to suffer severe effects if they do contract the illnesses. Accordingly, those groups are highlighted as populations of special concern.

Mortality and Morbidity

Commercially canned foods have rarely been linked to documented deaths in the United States. Botulism is the pathogen generally considered to be of most concern in regard to canned foods, but almost all recent cases of botulism caused by canned foods in this country have been linked to domestic home-canned goods. According to CDC, a total of six botulism deaths in 10 years were recorded as due to other food products (CDC 2010), none of them associated specifically with commercial canned foods. Scallan et al. (2011a) estimate 0-55 deaths per year from botulism poisoning.

According to outbreak surveillance data from CDC (2010), from 1998 to 2008 a total of 11 cases of foodborne botulism were attributed to canned foods; the data include illnesses from consumption of home-prepared products. In recent years, there has been only one recorded event in which commercially canned foods were linked to botulism. In 2007, four people got sick after consumption of a domestic brand of beans in the first recorded outbreak of botulism linked to commercially canned foods in 30 years (CDC 2007). As discussed above, foodborne illnesses are assumed to be under-reported, but Scallan et al. (2011a) estimated the under-reporting and under-diagnosis rate of botulism as relatively low (about a factor of 2).

A number of other natural and artificial toxicants may be present in canned foods, such as pesticide residues, mycotoxins, and chemicals that leach from packaging into the product. Available data are not sufficient to assess their effects on health. Reported cases of any of the agents in canned foods are also rare. In 1989, imported canned mushrooms from China were responsible for as many as four outbreaks of staphylococcal intoxication in different parts of the United States (CDC 1989). The outbreaks involved 102 people, 7 of whom required hospitalization, but no deaths were reported. In a 2003 outbreak involving 65 cases of salmonellosis, canned mushrooms were again implicated, but it was not clear whether the contamination occurred after cans were opened. In 2007, three people suffered scombroid intoxication from canned tuna (CDC 2010). The recent isolated outbreaks were documented as linked to domestic or imported canned products. There is no record of any death associated with commercially produced canned foods in the last 30 years.

To estimate mortality and morbidity attributable to canned foods, the committee used the data on deaths and illness described above, mortality and hospitalization rates from Scallan et al. (2011a) for each of the pathogens described above, and methods similar to those described above for leafy greens and shrimp. Specifically, the values presented in Table 4-1 were estimated at follows:

- *Low Estimate*. The committee assumed that no botulism cases attributed to commercially canned food occurred for the low estimate. The outbreak data described above were annualized and then adjusted by using the multipliers described in Scallan et al. (2011a) for under-reporting to obtain estimates of numbers of cases of illness associated with canned goods (by pathogen). Numbers of deaths, severe health effects, less severe health effects, and adverse quality-of-life health effects were then calculated according to the general method used for leafy greens.
- *Best Estimate*. The best estimate differs from the low estimate only in the attribution of botulism cases. For the best estimate, the committee assumed that 10% of botulism cases could be attributed to commercially canned foods.
- *High Estimate*. The high estimates were calculated by assuming that half the botulism cases in a year could be attributed to commercially canned foods and that the total incidence of the other foodborne illnesses described above (that is, 186 reported illnesses from canned foods) occurs in a single year. Those data were then adjusted by using the multipliers described in Scallan et al. (2011a,b) for under-reporting and under-diagnosis. Numbers of deaths, severe health effects, less severe health effects, and adverse quality-of-life health effects were then calculated according to the general method used for leafy greens.

Personal Controllability

As with the other food categories discussed, people theoretically have the ability to eliminate their personal risks of foodborne illnesses from commercially canned foods by avoiding the use of canned foods entirely. Given the prevalence of canned foods, that solution is not practical. Avoiding the use of any canned foods that show signs of spoilage—such as bulging, leaking, or dented cans—can help to minimize the chances of illnesses. Assuming that commercially canned foods are used just as frequently in home preparation of foods as in restaurant preparation, that about 27% of all meals consumed in the United States are prepared outside the home (Lin et al. 1998), and that 60-80% of consumers are aware of the precautions that should be taken with damaged commercially canned foods, about 45-60% of all cases of foodborne illnesses from commercially canned foods could be avoided through actions by individual consumers

Ability to Detect Adverse Health Effects

As noted above in connection with leafy greens and shrimp, the ability of an institution to detect adverse effects is problematic because illness has to be identified and then correctly attributed to a specific food product. Canned foods have a much longer shelf-life than leafy greens and shrimp and a well-documented production process and distribution system, so the possibility of attributing illness to them may be greater. However, canned foods will most likely be consumed over a longer period, so any illnesses associated with them may be seen as sporadic cases and be more difficult to identify and attribute to a source. Furthermore, because commercially canned foods are rarely associated with foodborne illnesses, individual cases are more likely to be attributed to other foods than to canned foods. The committee, however, notes that the ability to detect botulism poisoning caused by consumption of canned foods is quite high given that public-health officials must report a single case to CDC so that it can be investigated. Overall, however, other foodborne illnesses associated with canned foods are much more likely to occur than botulism poisoning, and they are more difficult to detect and attribute successfully to canned foods. On the basis of the under-reporting and under-diagnosis multipliers used in Scallan et al. (2011a), the committee estimated that about half the cases of botulism poisoning from commercially canned foods would be detected, and less than 3% of all other types of foodborne illnesses from canned foods would be detected.

Ability to Mitigate Adverse Health Effects

Industrial canning processes are typically designed to deliver products "commercially sterile" to minimize the presence of pathogenic organisms and to extend the shelf-life of foods. Commercial sterility is defined in 21 CFR 113.3 as the process in which heat is applied to render food free from "(a) microorganisms capable of reproducing in the food under normal non-refrigerated conditions of storage and distribution; and (b) viable microorganisms (including spores) of public health significance." Canned foods are typically produced by packing at near-boiling temperatures in tightly sealed containers. The hot-fill process eliminates most of the oxygen and creates a strictly anaerobic condition that minimizes oxidative reactions and inhibits aerobic microorganisms. At the same time, the initial high-temperature step kills all microbial vegetative cells. The tight seal of canned foods protects the product from external contamination during further processing and during distribution. The key step in the manufacture of canned foods is the treatment of recently sealed cans at temperatures above 100˚C for several minutes (Murano 2003). That heat treatment is typically conducted inside retorts, pressurized containers that use steam to attain temperatures as high as 121˚C (Potter and Hotchkiss 1995). After thermal processing, cans are normally cooled with water and stored until they enter distribution.

In addition, canned foods are tightly regulated (see, for example, 21 CFR 113, 108.25, and 108.35), including requirements for the proper operation and design of facilities that produce canned foods (Section 113). Section 113 defines two types of canned foods—acidified and low-acid foods—on the basis of the ability of *C. botulinum* to grow at a pH above 4.6. The regulations described in Section 113 apply mostly to low-acid foods. The rules in Section 113 include the characteristics of the equipment, controls, facilities, and product preparation that producers need to comply with. The regulations also require reporting any deviation in process characteristics of and continuous recordkeeping on every processed batch. As a result of that heightened oversight, canned foods have rarely been linked to cases of foodborne disease in the last few decades. The effectiveness of the steps to prevent contamination of canned foods can be seen in the very low rates of foodborne illnesses from canned foods shown in Table 4-1.

If a problem with canned foods does occur and is detected the ability of institutions to mitigate adverse effects through successful recall of a contaminated product before it is consumed is substantially higher than for the other two kinds of products considered in this case study, primarily because of the longer shelf-life of canned foods. The supply chain for canned foods is complex; a single manufacturing facility sometimes produces products that carry many different labels. The process, however, is well understood and relatively easily tracked; if a problem is detected and successfully traced to a particular facility, the potentially contaminated products can be readily identified in appropriate recall notices. For canned foods, the potentially contaminated product is likely to be widely distributed by the time a problem is detected, so the recall itself can be challenging; that is, all affected retail outlets and consumers must be made aware of and respond to the recall if it is to be completely effective.

After the 2007 Castleberry recall of tens of millions of cans because of potential *C. botulinum* contamination, a review by the North Carolina Department of Agriculture and Consumer Services found that the product remained on the shelves of 38% of the retail outlets that handled it (Seltzer et al. 2008). Most were smaller retail outlets, so this does not necessarily represent 38% of all recalled product, but it highlights a weakness in the effectiveness of recalls even for a well-understood supply chain and a product with a long shelf-life. Recalls are even less effective for product that has already been purchased by consumers. Patrick et al. (2007) report on the result of a random telephone survey conducted after a large-scale nationwide recall: only 45% of all adults were even aware of the recall. Again, depending on the severity of the foodborne illness caused by the canned food, a recall may come early enough for most of the product to still be in stores and not in individual consumers' homes, so that does not necessarily imply that 55% of the recalled product will be consumed. On the basis of those studies, the committee estimated that the ability of institutions to mitigate the adverse effects of foodborne illness caused by consumption of canned foods after a potential problem has been detected is about 50-75%.

USING THE RISK CHARACTERIZATION TO SUPPORT DECISION-MAKING

A review of Table 4-1 reveals a clear "ranking" of the three food categories in terms of the number of foodborne illnesses associated with each: domestic leafy greens appear to cause substantially more illnesses than do shrimp, and both appear to cause far more illnesses than do commercially canned foods. Other factors, however, also differ between the food categories and could be relevant to decision-makers and policy-makers. For example, people are estimated to have a higher degree of control over their own risk of contracting a foodborne illness from leafy greens than from shrimp. That information could be interpreted in several ways. It might suggest, for example, that efforts to improve awareness of how to prepare foods safely at home would do more to reduce the number of foodborne illnesses from leafy greens than to reduce the number from shrimp, whereas restaurant-focused efforts might be more effective in reducing the number of shrimp-related illnesses. Similarly, the higher ability to detect and mitigate risks associated with foods with well-understood and controlled distribution channels and longer shelf-lives, such as commercially canned foods, could suggest that supply-chain management would be

helpful for leafy greens or that efforts to reduce the number of foodborne illnesses from leafy greens should focus on the source because mitigation after exposure is more difficult.

As discussed in Chapter 2, although a risk ranking is not necessarily directly useful for a decision-maker, it can be a useful first step in making targeting or resource-allocation decisions. For example, if FDA were considering a high-level decision about whether to focus newly available inspection resources on leafy greens, shrimp, or commercially canned foods, the risk characterization developed here would provide information on the current levels of adverse health effects associated with each category and on where the maximum potential for risk reduction exists. Deciding how to allocate the resources would require additional analysis and more detailed understanding of how the resources would be used. If resources were available to inspect a particular number of agricultural suppliers in the field or some number of import locations or some number of canned-food production facilities, additional risk characterizations (attribute tables) would need to be developed that describe the likely outcomes of the increased inspections. Developing those risk characterizations would require consideration of how much of the particular food type could be inspected and the effectiveness of the inspections in reducing contamination in addition to all the factors considered above about numbers, types, and severity of foodborne illnesses. The differences between the attribute tables developed above and the attribute tables describing the public-health consequences of each food category in the enhanced inspection would be a measure of the relative benefits of different resource allocations.

REFERENCES

Bender, J.B., C.W. Hedberg, J.M. Besser, D.J. Boxrud, K.L. MacDonald, and M.T. Osterholm. 1997. Surveillance by molecular subtype for Escherichia coli O157:H7 infections in Minnesota by molecular subtyping. N. Engl. J. Med. 337(6):388-394.

Beuchat, L.R. 1996. Pathogenic microorganisms associated with fresh produce. J. Food Protect. 59(2):204-216.

Calvin, L. 2007. Outbreak Linked to Spinach Forces Reassessment of Food Safety Practices. Amber Waves (June):25-31 [online]. Available: http://www.ers.usda.gov/AmberWaves/June07/PDF/Spinach.pdf [accessed Oct. 26, 2010].

Calvin, L., H.H. Jensen, and J. Liang. 2009. The economics of food safety: The 2006 foodborne illness outbreak linked to spinach. Pp. 399-418 in Microbial Safety of Fresh Produce, X. Fan, B.A. Niemira, C.J. Doona, F.E. Feeherry, and R.B. Gravani, eds. Ames, IA: IFT Press.

CDC (U.S. Centers for Disease Control and Prevention). 1989. Epidemiologic notes and reports multiple outbreaks of staphylococcal food poisoning caused by canned mushrooms. MMWR 38(24):417-418 [online]. Available: http://www.cdc.gov/mmwr/preview/mmwrhtml/00001410.htm [accessed Oct. 26, 2010].

CDC (U.S. Centers for Disease Control and Prevention). 2007. Botulism associated with commercially canned chili sauce--Texas and Indiana, July 2007. MMWR 56:1-3 [online]. Available: http://www.cdc.gov/mmwr/preview/mmwrhtml/mm56d730a1.htm [accessed Oct. 27, 2010].

CDC (U.S. Centers for Disease Control and Prevention). 2010. Outbreak Net, Foodborne Outbreak Online Database. Available: http://wwwn.cdc.gov/foodborneoutbreaks/Default.aspx [accessed Oct. 27, 2010].

FAO (Food and Agricultural Organization of United Nations). 2010. Shrimp Market Report. GLOBEFISH, Fisheries Department, Food and Agricultural Organization of United Nations. May 2010 [online]. Available: http://www.globefish.org/shrimp-market-reports.html [accessed Oct. 27, 2010].

FDA (U.S. Food and Drug Administration). 2010a. Glossary of import term: Line (Line item). Subchapter 6.7.27 in Investigations Operations Manual. Inspections, Compliance, Enforcement, and Criminal Investigations, U.S. Food and Drug Administration [online]. Available: http://www.fda.gov/ICECI/Inspections/IOM/ucm063172.htm [accessed Mar. 21, 2011].

FDA (U. S. Food and Drug Administration). 2010b. Archive for Recalls, Market Withdrawals & Safety Alerts. U. S. Food and Drug Administration [online]. Available: http://www.fda.gov/Safety/Recalls/ArchiveRecalls/default.htm [accessed Oct. 27, 2010].

The Fish Site. 2008. U.S. Shrimp Market Report - April 2008 [online]. Available: http://www.thefishsite.com/articles/431/us-shrimp-market-report-april-2008 [accessed July 2010].

FSIS (Food Safety and Inspection Service). 2006. Safe Food Handling: Cleanliness Helps Prevent Foodborne Illness. Fact Sheet. Food Safety and Inspection Service, U.S. Department of Agriculture [online].

Available: http://www.fsis.usda.gov/factsheets/Cleanliness_Helps_Prevent_Foodborne_Illness/index.asp [accessed Mar. 22, 2011].

GAO (U.S. Government Accountability Office). 2008. Federal Oversight of Food Safety: FDA's Food Protection Plan Proposes Positive First Steps, but Capacity to Carry Them Out Is Critical. GAO-08-435T. Washington, DC: U.S. Government Accountability Office [online]. Available: http://www.gao.gov/new.items/d08435t.pdf [accessed Mar. 21, 2011].

GAO (U.S. General Accountability Office). 2010. FDA Could Strengthen Oversight of Imported Food by Improving Enforcement and Seeking Additional Authorities. GAO-10-699T. Washington, DC: GAO [online]. Available: http://www.gao.gov/new.items/d10699t.pdf [accessed Oct.27, 2010].

Hallman, W.K, C.L. Cuite, J.E. Dellava, M.L. Nucci, and S.C. Condry. 2009. Public response to the 2006 recall of contaminated spinach. Pp. 351-368 in Microbial Safety of Fresh Produce, X. Fan, B.A. Niemira, C.J. Doona, F.E. Feeherry, and R.B. Gravani, eds. Ames, IA: IFT Press.

Hoffmann, S., P. Fischbeck, A. Krupnick, and M. McWilliams. 2007. Using expert elicitation to link foodborne illnesses in the United States to foods. J. Food. Prot. 70(5):1220-1229.

IOM (Institute of Medicine). 2007. Seafood Choices: Balancing Benefits and Risks. Washington, DC: The National Academies Press.

IOM/NRC (Institute of Medicine and National Research Council). 1998. Ensuring Safe Food: From Production to Consumption. Washington, DC: National Academy Press.

IOM/NRC (Institute of Medicine and National Research Council). 2003. Scientific Criteria to Ensure Safe Food. Washington, DC: The National Academies Press.

IOM/NRC (Institute of Medicine and National Research Council). 2010. Enhancing Food Safety: The Role of the Food and Drug Administration. Washington, DC: The National Academies Press.

Koonse, B., W. Burkhardt, S. Chirtel, and G. Hoskin. 2005. Salmonella and sanitary quality of aquacultured shrimp. J. Food Protect. 68(12):2527-2532.

Lin, B.H., J. Guthrie, and E. Frazao. 1998. Popularity of dinning out present barrier to dietary improvements. Food Rev. 21(2):2-10 [online]. Available: http://www.ers.usda.gov/publications/foodreview/may1998/may98a.pdf [accessed Mar. 22, 2011].

Mandrell, R.E. 2009. Enteric human pathogens associated with fresh produce: Sources, transport and ecology. Pp. 5-42 in Microbial Safety of Fresh Produce, X. Fan, B.A. Niemira, C.J. Doona, F.E. Feeherry, and R.B. Gravani, eds. Ames, IA: IFT Press.

Mead, P.S., L. Slutsker, V. Dietz, L.F. McCaig, J.S. Bresee, C. Shapiro, P.M. Griffin, and R.V. Tauxe. 1999. Food-related illness and death in the United States. Emerg. Infect. Dis. 5(5):607-625.

Murano, P.S. 2003. Understanding Food Science and Technology. Belmont, CA: Thomson/Wadsworth.

NCI (National Cancer Institute). 2010. Usual Dietary Intake: Food Intakes, U.S. Population 2001-04. Risk Factor Monitoring and Methods. National Cancer Institute, U.S. National Institute of Health [online]. Available: http://riskfactor.cancer.gov/diet/usualintakes/pop/ [accessed Oct. 26, 2010].

NOAA (National Oceanic Atmospheric Administration). 2010. Fisheries of the United States – 2009, E.S. Pritchard, ed. National Marine Fisheries Service, Office of Science and Technology, National Oceanic Atmospheric Administration, Silver Spring, MD [online]. Available: http://www.st.nmfs.noaa.gov/st1/fus/fus09/fus_2009.pdf [accessed June 18, 2010].

Patrick, M.E., P.M. Griffin, A.C. Voetsch, and P.S. Mead. 2007. Effectiveness of recall notification: Community response to a nationwide recall of hot dogs and deli meats. J. Food Prot. 70(10):2373-2376.

Potter, N.N., and J.H. Hotchkiss. 1995. Food Science, 5th Ed. New York: Chapman and Hall.

Rounds, J.M., C.W. Hedberg, S. Meyer, D.J. Boxrud, and K.E. Smith. 2010. Salmonella enterica pulsed-field gel electrophoresis clusters, Minnesota, USA, 2001–2007. Emerg. Infect. Dis. 16(11):1678-1685.

Scallan, E., R.M. Hoekstra, F.J. Angulo, R.V. Tauxe, M.A. Widdowson, S.L. Roy, J.L. Jones, and P.M. Griffin. 2011a. Foodborne illness acquired in the United States--major pathogens. Emerg. Infect. Dis. 17(1):7-15.

Scallan, E., P.M. Griffin, F.J. Angulo, R.V. Tauxe , and R.M. Hoekstra. 2011b. Foodborne illness acquired in the United States--unspecified agents. Emerg. Infect. Dis. 17(1):16-22.

Seltzer, J., J. Rush, and J. Kinsey. 2008. Castleberry's: 2007 Botulism Recall; A Case Study by the Food Industry Center, August 2008. Food Industry Center, University of Minnesota [online]. Available: http://ageconsearch.umn.edu/bitstream/46055/2/Castelberry%20Study-%20Final.pdf [accessed Oct. 28, 2010].

Tooze, J.A., D Midthune, K.W. Dodd, L.S. Freedman, S.M. Krebs-Smith, A.F. Subar, P.M. Guenther, R.J. Carrol, and V. Kipnis. 2006. A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. J. Am. Diet. Assoc. 106(10):1575-1587.

Tsutsumi, H. 2007. A Food Predictive Modeling and Risk Estimation for Imported Shrimp into Japan and the United States. MS Thesis, University of Minnesota, St. Paul, MN. March 2007.

USDA (U.S. Department of Agriculture). 2009. Canned Fruits, Canned Vegetables, Canned Meats. Food Availability (Per Capita) Data System. Economic Research Services, U.S. Department of Agriculture [online]. Available: http://www.ers.usda.gov/data/foodconsumption/FoodAvailspreadsheets.htm#vegcan [accessed June 18, 2010].

USDA (U.S. Department of Agriculture). 2010. U.S. per Capita Food Availability: Custom Report-Individual Vegetable-Spinach. Economic Research Services, U.S. Department of Agriculture [online]. Available: http://www.ers.usda.gov/Data/FoodConsumption/app/reports/displayCommodities.aspx?reportName=Individual+vegetable&id=24#startForm [accessed Oct.26, 2010].

U. S. Census Bureau. 2010. Population Estimates, Population Characteristics, National Sex and Age [online]. Available: http://www.census.gov/popest/national/asrh/NC-EST2009-sa.html [accessed Oct. 26, 2010].

Wan Norhana, M.N., S.E. Poole, H.C. Deeth, and G.A. Dykes. 2009. Prevalence, persistence and control of *Salmonella* and *Listeria* in shrimp and shrimp products: A review. Food Control 21(4):343-361.

5

Case Study of a Strategic-Investment Decision

This chapter describes a case study that uses the committee's framework to characterize the public-health consequences associated with two medical devices under current and enhanced postmarket-surveillance programs. The decision of whether to implement the enhanced postmarket-surveillance system is an example of a strategic-investment decision. The case study was selected because it is relevant to a scenario provided to the committee by the Food and Drug Administration (FDA) and there was relevant expertise on the committee. The data used were gleaned from publicly available Web sites or publications or were provided by FDA. The committee did not conduct exhaustive literature searches or reviews, and all information is illustrative. The case study simply provides an illustration of how the committee's framework might be used for a strategic-investment decision.

FRAMING THE ISSUE: MEDICAL DEVICES AND POSTMARKET SURVEILLANCE

The FDA Center for Devices and Radiological Health regulates an estimated 220,000 devices, reviews around 3,500-4,000 new products each year, and monitors some 1,000 product recalls each year. Items as varied as tongue depressors, tubing, and pacemakers are under the purview of this FDA center. At the approval stage, medical devices are labeled as Class I, II, or III on the basis of the control needed to ensure the safety and effectiveness of the devices; Class III products are life-supporting and life-sustaining devices. For recalls, the order of severity is reversed. Class I recalls involve dangerous or defective products that predictably could cause serious health problems or death, such as a defective artificial heart valve; Class II recalls involve products that could cause temporary health problems; and Class III recalls involve products that are unlikely to cause any adverse health effect. In 2009, FDA reported that there were 160 Class I recalls, over 500 Class II recalls, and 172 Class III recalls (FDA 2010a). Not all the recalled products were in clinical use, so the recalls may have prevented additional future exposures.

Many recalls of implanted medical devices are based partly on postmarket surveillance data. A universal, one-size-fits-all system for medical-device reporting and postmarket surveillance has not yet reached the level of success that FDA would like. The existing FDA medical-device reporting systems—Manufacturer and User Facility Device Experience (MAUDE) and Medical Device Reporting (MDR)—contain a great deal of information that cannot be reliably analyzed, and the agency is studying new approaches to postmarket surveillance under the Sentinel Initiative.

Under MDR requirements, manufacturers of medical devices are required to report deaths and serious injuries caused by malfunctions of medical devices to FDA. User facilities are required to report serious injuries associated with medical devices to their manufacturers and to report deaths to both the manufacturers and FDA (FDA 2009a). Data from 1991 through 1996 are in the MDR database, and data

from various sources from 1991 to the present are in the MAUDE database (manufacturer and user-facility reports since 1996 are included in MAUDE) (FDA 2010b). Although the systems include useful information on the number and types of adverse events associated with medical devices, they do not include clear information on the number of people who use the devices, and that makes it difficult to estimate the *rate* of such adverse events and therefore difficult to detect changes in the rate.

Medical-device registries have some appeal, but it is probably not possible to implement them for all medical implants that are in general clinical use. There is a device registry in place for ventricular assist devices (VADs): INTERMACS® (Interagency Registry for Mechanically Assisted Circulatory Support) is a national registry for patients who are receiving mechanical circulatory-support device therapy to treat advanced heart failure. This registry was devised as a joint effort of the National Heart, Lung, and Blood Institute, the Centers for Medicare and Medicaid Services, and FDA and was formed to analyze clinical outcomes, device durability, adverse-event rates, and costs. Analysis of the data collected is expected to facilitate improved patient evaluation and management and aid in better device design and development. Registry results are also expected to influence future research and system design and facilitate appropriate regulation of VAD implants. About 85% of all VAD implants are enrolled in the registry (D.C. Naftel, University of Alabama at Birmingham, personal. commun., March 7, 2011).

DECISION CONTEXT FOR THE CASE STUDY

This case study focuses on medical implants that make direct contact with tissues other than the skin. They constitute a pool of about 20,000 medical implants designed for various periods of use (months to years) and are manufactured by domestic and foreign companies. A paracorporeal infusion pump is defined as a medical implant for this study, and transdermal patches worn on the skin are not. The rationale for a focus on medical implants is partly the postmarket-surveillance scenario identified by FDA (see Appendix C) and partly the fact that patient compliance does not have to be considered as a factor in estimating the effects (the patient is by definition compliant with a medical implant because it can be modified only with surgery).

The strategic-investment decision considered here is whether FDA keeps the reporting system as it is or invests in enhanced postmarket surveillance of two specific medical devices—artificial knees and VADs. The former is an established technology that is used extensively across the country and has a long history of clinical use in a large patient population. The latter is an emerging technology in limited use and, as noted above, the subject of a comprehensive registry (INTERMACS).

There are many ways that an enhanced postmarket-surveillance system could be designed, but the goal should be to find types and patterns of unexpected adverse events. Such patterns may point to problems in design, implantation processes, clinical interventions, or manufacturing variances; early detection of such problems should lead to improvements. The information gathered in a postmarket-surveillance system would be of value not only to FDA but to the medical-device industry and to individual patients. FDA's New Molecular Entity Postmarketing Safety Evaluation Pilot Program to evaluate accumulated information within some specified period after a drug is approved for marketing is an example of such a reporting system (FDA 2009b). Systems currently under development include FDA's Sentinel Initiative and MDEpiNet programs, which are intended to provide new surveillance capabilities for device-related adverse events (FDA 2011a,b). Better information on the effects of medical implants on individual patients—in particular, insights from patient social networks as noted below—could also be useful for patients and physicians engaged in shared decision-making (Charles et al. 1997).

The committee defined the enhanced surveillance system for this case study as one that would require manufacturers and hospitals—which are more able to determine the number of medical

implants—to report both numbers and rates of device-related deaths and unexpected adverse effects and to report marked changes in adverse-effect rates. With modern information-technology capabilities, it may be feasible now to develop systems to flag and report effect-rate changes. The enhanced system would also include direct and indirect efforts to gain more information directly from patients by, for example, using social networks for rapid patient feedback on quality of life and health issues related to the implants. Finally, the hypothesized postmarket-surveillance system would include not only additional data collection and tracking but analysis of adverse-event data and, when appropriate, would incorporate lessons from the surveillance data into patient-selection guidelines and recommendations for postimplantation care.

CHARACTERIZING THE PUBLIC-HEALTH CONSEQUENCES

This section summarizes the use of the risk attributes to characterize the public-health consequences associated with the current postmarket surveillance and the hypothetical enhanced surveillance of artificial knees and VADs. The committee estimated the values for each of the risk attributes in the current system by using scientific literature and subjective judgments based on available information (given its limitations). For the enhanced system, the values were developed on the basis of several overarching assumptions: that enhanced surveillance would lead to better understanding of the risks and the risk factors associated with each device and that action would be taken according to that information, as appropriate, to improve patient outcomes through improved patient selection and postimplantation care and follow up. The committee also assumed that better tracking of adverse events improves the ability to prevent adverse effects, to detect them if they occur, and perhaps to mitigate them. Table 5-1 summarizes the comparison.

Several challenges that are unique to evaluation of postmarket surveillance of implanted medical devices arose in this case study. First was the issue of data availability and the current state of information-tracking and reporting responsibilities. Although the most important and most useful information to have is information on rates of unexpected adverse events, those data are not collected or reported. When a person who has an implant (or more than one implant) dies, the death is reported, but the death may or may not be related to the implant. The lack of available data leads to substantial uncertainty in the estimates of the number of adverse health effects associated with the devices, as shown in Table 5-1. Second was the difficulty of developing estimates of the effects under an enhanced postmarket-surveillance system without a detailed definition of what that system would look like. Many ways of designing such a system could be envisioned, and for this case study the committee assumed only that improved information on adverse-event rates would somehow be obtained.

Artificial Knees

Artificial knees provide increased mobility and decreased pain and are generally implanted in adult patients with relatively good health. The mortality risk due to the knee replacement is primarily in the 90-day postoperative period. In terms of quality of life, people can resume hobbies or work activities that were restricted because of pain. Moss et al. (1991) contains 1988 data on selected medical-device implants in the United States, including artificial joints and heart valves. Data are provided on type of implant, number of each implant type (such as two artificial joints), socioeconomic characteristics, and reasons for implantation. Additional information on artificial knees can be found in AAOS (2009) and Palmer and Cross (2010).

TABLE 5-1 Risk Attributes for Strategic-Investment Decision for Medical Devices

		Decision Options			
		Artificial Knees		Ventricular Assist Devices	
Attribute	Metric	Current System	Enhanced System	Current System	Enhanced System
Exposed population	Number who use the product or resource in a year	5 million (2-8 million)	5 million (2-8 million)	3,000 (1,000-5,000)	3,000[a] (1,000-5,000)
	Populations of concern	Most of the exposed population are elderly; women are twice as likely as men to receive implants; other populations with higher than average exposure are people suffering from arthritis		The exposed population consists entirely of people with advanced heart failure; men are 4 times as likely as women to require the implants	
Mortality	Number of deaths per year	6,000 (3,000-150,000)	Reduce the uncertainty (the range) by about 85%; for example, the range might be 4,000-26,000	300 (200-400)	Reduce the uncertainty (the range) by about 35%; for example, the range might be 220-350
Morbidity	Number experiencing severe adverse health effects per year	80,000 (15,000-130,000)	Reduce the uncertainty (the range) by about 30%; for example, the range might be 20,000-100,000	410 (240-510)	Reduce the uncertainty (the range) by about 40%; for example, the range might be 300-460.
	Number experiencing less severe adverse health effects per year	1.25 million (1-1.5 million)	No change	1,500 (1,000-2,500)	No change
	Number per year experiencing adverse health effects that affects only quality of life	No estimates developed			
Personal controllability	For operative and 90-day postoperative risks	100% of patients have the ability to avoid or reduce the risks associated with implantation of an artificial knee		Less than 40% of patients have the ability to avoid or reduce the risks of VAD implantation.	
	For risks associated with living with the implants	5-10% of problems with implanted artificial knees can be reduced or managed through personal action by patients	10-20% of problems with implanted artificial knees can be reduced or managed through personal action by patients	Less than 5% of patients have the ability to control or reduce the risks associated with an implanted VAD after surgery	
Ability to detect adverse health effects	Ability of informed institution to detect population-level effects associated with product being evaluated	10-25% of adverse health effects caused by artificial knees could be detected and successfully attributed	25-75% of adverse health effects caused by artificial knees could be detected and successfully attributed	About 90% of adverse effects caused by VADs would be detected and correctly attributed	
Ability to mitigate adverse health effects	Probability that an informed institution will be able to reduce or mitigate any adverse health effects associated with the specific product being evaluated if such a problem is known to exist	80%	90%	80%	90%

[a]If the sole VAD on the market is recalled, there would be no more implants.

Exposed population

The committee used data from several sources to estimate the number of people who are currently living with at least one artificial knee. Because artificial knees carry perioperative risks and some continuing risks associated with living with the device, the committee considered data on the annual number of implants and on the total number of people who live with implants as relevant. Recent data indicate that about 581,000 knee replacements are performed each year (AAOS 2009). By 1988, the *total* number of people who had knee implants in the United States was 521,000 (Moss et al. 1991); however, no recent data were available on the total number of people living with implants. The committee assumed that artificial knees are in use for about 20 years and that some patients die of causes unrelated to the implants and some require revisions. The committee made a direct judgment that the number of people receiving or living with one or two artificial knees in a given year is about 5 million but could range from 2 million to 8 million. In the terminology of Chapter 2, 5 million is the best estimate of the size of the exposed population, 2 million is the low estimate (defined as about equal to the 5th percentile of a probability distribution of the exposed population), and 8 million is the high estimate (defined as about equal to the 95th percentile).

For the enhanced surveillance system, the committee estimated that there would be no change in the size of the exposed population. Those who currently have knee replacements retain those implants and remain part of the population exposed to the long-term risks associated with the devices. Although new information from the enhanced system might lead to changes in guidance on patient selection for implant operations, the committee concluded that the enhanced system would be unlikely to change the aggregate numbers receiving knee replacements.

Knee-replacement recipients are mainly older adults who have mobility limitations, but they also include teenagers who have arthritis. Roughly twice as many women receive knee replacements as men (among Medicare-funded operations), according to data from Katz et al. (1996). The committee assumed that the sex ratio would not change in the enhanced surveillance system.

Mortality and Morbidity

Most deaths of knee-replacement patients occur in the 90-day postoperative period, although infections or other adverse effects can occur in the later years of living with the device. Complications that may arise after total knee-replacement surgery include blood clots, infection, patellofemoral complications, neurovascular complications, fractures around the prosthetic, loosening of the prosthetic, and excess scar tissue that can cause restriction of knee movement. Most of the adverse effects occur in no more than 1% of patients; patellofemoral complications are the most common reason for reoperation. Loosening has the highest incidence (5-10% of patients 0-15 years after initial surgery). There have been improvements in devices, and those with more recent implants may face lower risks than those living with older implants.

There are a variety of mortality estimates in the literature (see Table 5-2 for selected values), from which the committee estimated 0.5% 90-day postoperative mortality rate.

For *mortality* under the current system, the committee calculated estimates as follows:

- *Low Estimate*. The low estimate is the product of 90-day postoperative mortality rate (0.5%) and the number of operations performed each year (600,000), which yields 3,000 deaths each year.
- *Best Estimate*. The best estimate is 2 times the low estimate and includes both the 90-day postoperative deaths and an assumption that an equivalent number of device-related deaths occurs in the population living with artificial knees.
- *High Estimate*. For the high estimate, the committee estimated the number of artificial-knee recipients in the United States who would die in a year and attributed all those deaths to the implants. In the general U.S. population, eight of 1,000 people die each year from any cause; older people have higher

death rates, ranging from about 9 per 1,000 for people 55-64 years old to 50 per 1,000 for people 75-84 years old (NCHS 2010). Because artificial-knee recipients are typically older adults, the committee used an estimate of a 3% annual death rate in the exposed population of 5 million, for an estimated 150,000 deaths per year. The high estimate accounts for the fact that a large exposed population of 5 million who live with the devices presents the potential for a large number of people to die from some sort of implant-related complications.

Potential *severe adverse effects* of artificial knees include the perioperative risk of serious complications and the potential for adverse effects after implantation that may lead to the need for surgical intervention, such as patellofemoral complications, arterial thrombosis, and loosening. Mahomed et al. (2005) found the following 90-day postoperative complications associated with knee replacements among Medicare claims: 4.7% readmissions, 1.8% wound infection, 1.4% pneumonia, 1% myocardial infarction, and 0.5% pulmonary embolism. In a 7-year study of 39,286 primary total knee arthroplasties, Paxton et al. (2010) found that 1.7% were revised by the date of the study, 0.7% were revised because of infection, and 0.3% were revised because of instability.

On the basis of that information, the committee's low estimate of the number of severe adverse health effects is 2.5% of the 600,000 implants per year (that is, about half the rate of readmissions among Medicare claims). The best estimate of 80,000 severe adverse effects is based on an assumption that about 5% of all patients who receive an implant in a year will have one of the severe complications identified by Mahomed et al. and about 1% of those living with an artificial knee will experience a severe adverse effect associated with that implant. Each of the potential severe adverse effects described generally occurs in less than 1% of the exposed population except loosening, which is expected to occur in 5-10% of patients from 0-15 years after surgery. The high estimate is that such effects would occur in 2% of the exposed population combined with the 5% of the population who receive an implant and suffer a severe complication.

For *less severe effects*, the committee assumed that the effects would include nerve injury and less serious infections that could be treated with antibiotics. The best estimate is 25% of the exposed population of 5 million, the low estimate is 20% of the exposed population, and the high estimate is 30% of the exposed population.

TABLE 5-2 Mortality Estimates

Description	Value	Reference
Overall mortality rate	<1%	Palmer and Cross 2010
In-hospital 30-day mortality rate	0.12%	UAMS 2010
30-day mortality rates	0.41-0.73%	Taylor et al. 1997
	0.36%	Gill et al. 2003
	0.21%	Parvizi et al. 2001
90-day mortality rate	0.46%	Gill et al. 2003
90-day postoperative mortality rate after first surgery; rate derived from Medicare claims	0.7%	Mahomed et al. 2005
90-day postoperative mortality after revision surgery; rate derived from Medicare claims	1.1%	Mahomed et al. 2005
Cumulative survival rate at 5.5 years	97.1%	Paxton et al. 2010

The committee did not attempt to estimate the number of patients who experience adverse health effects important enough to affect their quality of life but not important enough to require medical attention or otherwise meet the definition of "less severe adverse effects." Although some such effects may occur, estimating them is exceptionally difficult in this case because knee replacements are associated with increases in quality of life, such as greater mobility and less pain; this greatly confounds any attempt to monitor or estimate adverse effects on quality of life that do not require some medical attention.

As described above, the hypothetical enhanced postmarket-surveillance approach defined for this case study focuses on gaining better information on the number and rates of adverse events through enhanced reporting requirements and direct and indirect patient outreach. The committee also assumes that the information collected through the enhanced surveillance would enable the medical community to take actions—to improve patient-selection criteria and to develop improved monitoring and follow-up of higher-risk patients. For example, there might be guidance on getting more x-ray examinations and clinical inspections of knees in higher-risk patients before and after implantation that would allow earlier detection and correction or avoidance of current or future life-threatening problems.

Those two features lead to two effects. First, improved information on adverse-event rates will reduce the uncertainty in the annual number of those adverse events (for example, the difference between the high and low estimates shown in Table 5-1 would be smaller with the enhanced surveillance than with the current system). It is not possible, however, to specify exactly how the high and low estimates would change, only that the difference between them would be smaller. For the annual number of deaths, for example, the current estimate is that 3,000-150,000 people die each year because of an artificial-knee implant. With better information, that range might be reduced by about 85%. But the end points of the estimated range could be 4,000-26,000, 5,000-27,000, or some other range; there is no credible way to estimate the precise end points of the range. Accordingly, Table 5-1 contains a brief description of the expected reduction in the uncertainty in the numbers of deaths and other adverse effects associated with the enhanced surveillance system and an example of what a new range could be (that is, what a more accurate estimate would look like).

Second, the assumption that the enhanced postmarket-surveillance system would enable medical providers to improve patient selection, monitoring, and follow-up suggests that the overall rates and annual numbers of deaths and severe adverse effects will be reduced over time as the improvements take effect. Because artificial knees have been in use for many years, much of the potential improvement in patient selection and monitoring may have already been attained, so the committee estimated negligible change in the number of adverse effects that would occur because of the added surveillance of the implants, at least in the short term.

Personal Controllability

The primary methods by which people can eliminate or reduce their own personal risk of death or serious 90-day postoperative adverse effects are by declining to have the knee replacement and by careful selection of hospitals and physicians with low adverse-effect rates. Medical informed consent requires that patients be informed of the risks associated with a proposed treatment or procedure, so it is reasonable to assume that they make the choice to undergo knee replacement willingly and with knowledge of the risks that they face from the surgery. Artifical knees are not a life-saving technology, so the committee assumed that all patients who receive an implant in a given year have the ability to control their risks (that is, could choose not to have the implant). Enhanced postmarket surveillance would not change that aspect of personal controllability.

For patients who already have an artificial knee, the ability to control the risks associated with that implant are substantially less. People have no ability to reduce the chances of device failure; they have limited ability to reduce the consequences of such failures through basic health maintenance and appropriate postoperative physical therapy and by seeking prompt medical attention if problems develop.

However, it may be difficult for a patient to detect a problem or to associate adverse health effects with the implant when they do not directly involve the knee implant itself, such as bloodborne infections that result from a dental treatment and effects on ligaments or muscles in other parts of the body. The committee estimated that 5-10% of the problems that arise with existing implants could be managed or reduced by individual actions by the patient. The hypothesized enhanced postmarket-surveillance system is designed to provide more and better information to the patient population and to the medical community and includes an assumption of increased use of social media to reach those patient populations. The committee estimates that the resulting increase in patient awareness of the various problems that can occur and steps that they can take to minimize the problems will lead to an increase in the controllability of the effects.

Ability to Detect Adverse Health Effects

Artificial knees that have been implanted more than a few years are examined when a patient chooses to return to an orthopedist; examinations are recommended every 2-3 years. However, the patient may not detect some problems or might not see an orthopedist about a possible problem even if walking becomes more difficult. Problems with knee replacements can lead to other problems (such as in ligaments or muscles in other parts of the body) that may not be readily recognized as being linked to the knees by patients or physicians. On the basis of the required reporting data for medical implants, it appears that only the device manufacturers would have sufficient data to identify systematic or population-level adverse effects, and they rely on adverse-event reporting from the facilities. Given the relatively long chain of reporting with multiple opportunities to miss signals, the committee judged that relatively few systematic problems would be detected—perhaps 10-25%.

In the enhanced surveillance system, the ability to detect a problem would improve. The hypothesized enhanced surveillance system focuses specifically on obtaining better, more accurate estimates of the rates of adverse events, and changes in those rates—exactly the type of information needed by institutions to detect emerging risks. The enhanced system delivers more accurate rate information that allows identification of higher-risk patient populations and improved guidance in patient selection and monitoring, all of which increases the ability to detect systematic problems. For example, risks of long-term effects, such as unexpected toxicity from device materials, might be detectable by institutions in a surveillance system that could institute enhanced monitoring of at-risk patients. The committee estimated that with enhanced surveillance systems, the ability to detect problems would increase by a factor of 2-3 and estimated that 25-75% of such problems could be detected and successfully attributed.

Ability to Mitigate Adverse Health Effects

As noted in Chapter 1, medical devices are highly regulated FDA products. The Center for Devices and Radiological Health conducts premarket reviews and monitors the manufacturing processes and uses of its regulated products. Premarket and postmarket inspections of manufacturing facilities are conducted, and imports are reviewed to ensure compliance with FDA standards. If a systematic problem does occur with a device, institutions can (or can attempt to) examine all relevant implant patients once that problem is detected. The adverse effects can typically be mitigated for the individual patient, although such mitigation may on occasion involve a substantial medical intervention, including possible replacement of the implant. The primary challenge is to identify and examine the relevant patients—determining who has an implant that might be at risk and encouraging the person to visit an orthopedist. Some patients will have moved and be lost to follow-up; others may have changed health-care providers, and the new providers may or may not have sufficient information available to them to identify a patient

as being at risk. The committee estimated that about 80% of all patients with a device with identified problems could be located and examined and have their problems corrected.

In the enhanced surveillance system, the probability that an institution can mitigate a problem is improved. Again, the improvement is likely to come from the fact that more accurate data on adverse events would lead to more accurate identification of at-risk patient groups, and more targeted outreach to those patients could occur. The committee estimated that in the enhanced surveillance system, the ability to mitigate adverse effects increases to about 90%.

Ventricular Assist Devices

VADs are medical devices implanted in patients who have advanced heart failure and are ineligible for heart transplantation. They are electromechanical systems designed to assume the work of a patient's left ventricle, improve coronary arterial perfusion, and provide systemic blood flow to all tissues and organs. They allow the generally bed-ridden patients to resume moderate activities.[1] Many advances have been made over the years with VADs, and more information can be found in Carr et al. (2010).

This medical-device class was chosen for this case study because it is an emerging technology that recently received FDA approval and the number of VAD patients is relatively small. The technology is entering clinical use in stepwise fashion with careful clinical, industrial, and federal involvement. In many respects, this novel medical technology may serve as a model for the orderly introduction of other emerging high-risk technologies and for enhancing the value of postmarket-surveillance systems for FDA decision-makers.

Exposed Population

For this case study, the committee considered only people who have advanced heart failure, who are generally not eligible for cardiac transplantation, and who may have a permanent VAD implanted. About 20,000-30,000 people may be eligible for VADs, but their use is not yet widespread, so the actual number of people exposed is much smaller than the population of potential VAD recipients. Estimates of exposure were based on data from the INTERMACS registry, the increasing use of VADs, and expert judgment. In 2010, about 3,000 people were listed in the registry as having received a VAD. The number of people living with a VAD may be higher or lower. For example, the INTERMACS registry is not complete; it is estimated that it contains data on 85% of the patients who have implants, and it may contain people who had a VAD and later received a heart transplant or died. On the basis of those data, the committee estimated that 1,000-5,000 people are living with VADs, representing the low and high estimates, respectively. The best estimate is that about 3,000 people are living with the devices. As the use of VADs increases, the number of people living with them will also increase; the data above and in the table provide a snapshot of the current status. The enhanced surveillance system is not expected to change the number of people receiving VADs.

Of the 3,000 people, about 20% are women and 80% are men (INTERMACS 2010). In the enhanced surveillance system, the committee concluded, there would not be a change in the sex ratio.

Mortality and Morbidity

To estimate *mortality* attributable to VADs, the committee considered both perioperative mortality and mortality occurring while people were living with the devices. The perioperative mortality

[1]Although VADs are used for shorter periods in patients who are to undergo cardiac transplantation, that use is not considered here.

rate for this population is about 5%; of the roughly 1,050 patients who receive VADs each year, about 50 would be expected to die (Kirklin et al. 2010). The annual mortality rate for patients who have VADs is about 10%; of about 3,000 patients who have VADs, about 300 would be expected to die each year. Using expert judgment and the data described, the committee estimated that 200-400 people will die each year in connection with use of VADs. The best estimate of the number of deaths is 300 (10% of the 3,000 people who have VADs). The committee notes that a VAD is a life-saving medical device: the mortality rate is reported to be about 80% for heart-failure patients who do not get VADs (Terracciano et al. 2010). However, for the purposes of this case study, the committee emphasizes that we are not providing an estimate of the benefit of having VAD technology available but rather are focused on the benefits of conducting enhanced postmarket surveillance of those who receive VADs.

To estimate the number of *severe adverse health effects*, the committee reviewed data on replacement and disabling stroke rates in VAD patients; stroke is a relatively common effect that meets the definition of *severe adverse effects* provided in Chapter 2. In clinical-trial patients, 16% had a baseline history of stroke, and 46% were free from disabling stroke and reoperation 2 years after implantation of a VAD (Slaughter et al. 2009). The same trial data indicate that there is 0.13 stroke per patient-year within 2 years after receiving a continuous-flow VAD implant and a 0.22 stroke per patient-year for pulsatile-flow VAD. The trial data also showed that 10% of devices were replaced at 2 years (that is, a 5% probability per year).

Using the probabilities for disabling stroke and for replacement, the committee estimated that in the current system a median of 410 people experience severe adverse health effects annually:

- 150 have to have VAD replacement within a year (5% of 3,000).
- 260 have strokes that are debilitating after receiving a VAD (0.13 stroke per patient-year for the 2,000 patients assumed to be within 2 years of initial replacement).

Using a VAD replacement rate of 3% and a disabling stroke rate of 5%, the committee estimated the 5th percentile to be 240, and using a VAD replacement rate of 7% and a disabling stroke rate of 10%, the committee estimated the 95th percentile to be 510.

Less severe adverse health effects that could follow VAD implantation include effects that would follow hospital release, such as drive-line infection, arrhythmia, renal or hepatic effects, thrombosis, and faulty battery or connections. In the current system, the committee assumed that 50% of those living with VADs experienced such effects, or 1,500 people experiencing less severe adverse health effects each year. The 5th percentile was judged to be 1,000, and the 95th percentile was judged to be 2,500.

The committee did not estimate the number of people who might experience adverse effects on quality of life because of VADs. As suggested above, VAD implantation is used to extend and improve the quality of life of those in advanced heart failure. Quality of life, for example, can improve from being bed-ridden to being capable of moderate exercise (Slaughter et al. 2009). Those benefits would obscure any adverse effects of the VAD that do not require medical treatment.

As described in the section on mortality and morbidity associated with artificial knees, the effects of improved postmarket surveillance of VAD patients would be to reduce the uncertainty about the rates and numbers of deaths and other adverse health effects and, if that information is acted on, potentially to reduce the number of such effects. As for artificial knees, there is no credible way to estimate the precise end points of the ranges of those numbers before implementing the surveillance program. Accordingly, Table 5-1 contains a brief description of the expected reduction in the uncertainty in the number of deaths and other adverse effects associated with the enhanced surveillance system rather than precise numbers. Because the current system of postmarket surveillance of VADs is already quite comprehensive and the patient population is small, the expected benefits of the enhanced system compared with the current system are smaller than in the case of artificial knees, for which the enhanced system is much more rigorous than the current one.

Personal Controllability

In contrast with artificial knees, VADs are life-saving medical implants used as a destination therapy for patients who have congestive heart failure and are ineligible for transplant. Those for whom the device is recommended are likely to see it as their option of last resort—the only opportunity to extend their lives and to improve their quality of life for the time that remains. Although they technically have the same option to decline surgery as do artificial-knee recipients, that is not likely to be viewed by most as a real option. No data are available on the percentage of patients who are offered a VAD and decline it; the committee estimated that less than 40% of patients in a situation where a VAD would be offered would consider it to be optional and could be considered to have the ability to control their personal risks.

For patients who already have a VAD, the ability to control the risks associated with the implant is substantially lower. People have no ability to reduce the chances of device failure, and VAD patients tend to be older and in poor health at the time of implant, further reducing their ability to exercise personal control over their VAD-related outcomes. The committee estimates that less than 5% of the problems that arise with existing implants could be managed or reduced by individual actions by the patient. Because of characteristics of the patient population, the hypothesized enhanced postmarket-surveillance system is not expected to change the personal controllability of postsurgical VAD risks.

Ability to Detect Adverse Health Effects

Medical professionals follow VAD patients closely, and the goal of the INTERMACS registry is to follow all VAD patients and identify signs of potential problems. Because of that close observation of VAD patients and the use of the INTERMACS registry, the committee concluded that the ability to detect systematic occurrence of adverse effects at the institutional level in either the current system or the enhanced system would be relatively high, that is, 90% or more of potential problems with a VAD would be detected and successfully attributed.

Ability to Mitigate Adverse Health Effects

As noted above, medical devices are highly regulated, and VADs are monitored carefully. However, VAD patients are generally in advanced heart failure and are more susceptible to surgical risks and to adverse events after implantation than are recipients of artificial knees. Although most complications of VADs do not require explant, some situations may require replacement or removal of a VAD, and that might not be possible in some cases because of the overall health of the patient. The committee assumed that in the current system there is an 80% chance that an institution will be able to mitigate any problems that are found that are directly related to the VAD. The committee assumed in the enhanced surveillance system that the probability is improved to 90% because problems may be able to be alleviated quicker with greater surveillance and quicker communication to clinicians.

USING THE RISK CHARACTERIZATION TO SUPPORT DECISION-MAKING

Table 5-1 highlights the public-health consequences associated with two devices under the current and enhanced postmarket-surveillance systems. The differences indicate the benefits of enhanced surveillance. Specifically, the direct public-health benefits of the hypothesized enhanced postmarket-surveillance system compared with the current one are (1) a reduction in the uncertainty about the number of adverse effects attributed to the devices because of better tracking and reporting of adverse effects and their causes, (2) an increase in the ability of informed institutions to detect systemic problems that may be

occurring with the devices because of the increased reporting requirements, and (3) a slight increase in the ability to mitigate adverse effects, again arising from the requirement for better reporting and tracking of problems and the assumption that the resulting improvement in understanding will lead to identification of ways to reduce adverse outcomes. For artificial knees, the enhanced surveillance system is also expected to increase the ability of individual patients to manage and reduce their own personal risks because of increased focus on patient outreach and information-sharing.

As in all the case studies, the differences in public-health consequences between the current and enhanced surveillance is only one of several factors that would need to be considered in deciding whether to make an investment in enhanced surveillance. Other factors that might be relevant include the costs and feasibility of implementing and validating an enhanced system, the ability to use the newly acquired data from surveillance to make decisions, and the usefulness of the improved information to the patients themselves in providing better information on patient-specific risks and benefits associated with elective devices.

This case study looked at a relatively simple comparison of "invest" or "do not invest" in an enhanced postmarket-surveillance system, but it raised some additional factors that could be of interest for more complex strategic-investment decisions or other types of decisions related to medical devices. For example, if the decision were to determine which of a variety of surveillance approaches should be pursued, a similar evaluation could be conducted. Rather than characterizing only two options, however, each surveillance approach would have to be defined and evaluated separately. If the decision were to determine on which products a new enhanced surveillance system should focus (that is, a targeting decision for a strategic investment), it may be useful to consider factors beyond the direct public-health benefits of the enhanced surveillance relative to the current system. In particular, applying an enhanced surveillance system to a product for which relevant, high-fidelity data are already being collected by another organization (for example, the INTERMACS registry for VADs described above) could provide a unique opportunity to compare the findings of the new enhanced surveillance system each year with an independent established dataset. That would provide a method for continuous improvement of the FDA surveillance system.

Finally, the committee notes that during the development of this case study, several issues related to medical devices arose that would probably be relevant for other device-related decisions. They include the speed with which health outcomes can be improved if a problem is detected, potentially measured as time between detection and correction; sustained health benefits of a medical implant and the performance of alternatives to the implant, which would be of particular relevance if FDA were evaluating decisions that could change the availability of the implant for potential recipients; and time-dependent projections of levels of exposure and health effects, especially for new products or ones whose use is growing or shrinking.

REFERENCES

AAOS (American Academy of Orthopaedic Surgeons). 2009. Total Knee Replacement. American Academy of Orthopaedic Surgeons [online]. Available: http://orthoinfo.aaos.org/topic.cfm?topic=A00389 [accessed Oct. 26, 2010].

Carr, C.M., J. Jacob, S.J. Park, B.L. Karon, E.E. Williamson, and P.A. Araoz. 2010. CT of left ventricular assist devices. RadioGraphics 30(2):429–444.

Charles, C., A. Gafni, and T. Whelan. 1997. Shared decision-making in the medical encounter: What does it mean? (or it takes at least two to tango). Soc Sci Med. 44(5):681-692.

FDA (U.S. Food and Drug Administration). 2006. Guidance for Industry and FDA Staff: Postmarket Surveillance under Section 522 of the Federal Food, Drug and Cosmetic Act. Center for Devices and Radiological Health, U.S. Food and Drug Administration [online]. Available: http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm072517.htm [accessed Nov.8, 2010].

FDA (U.S. Food and Drug Administration). 2009a. Reporting Adverse Events (Medical Devices). U.S. Food and Drug Administration [online]. Available: http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/PostmarketRequirements/%20ReportingAdverseEvents/default.htm#link [accessed Nov. 8, 2010].

FDA (U.S. Food and Drug Administration). 2009b. New Molecular Entity Postmarketing Safety Evaluation Pilot Program Final Report. U.S. Food and Drug Administration [online]. Available: http://www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/ucm185252.htm [accessed Oct. 26, 2010].

FDA (U.S. Food and Drug Administration). 2010a. FDA 101: Product Recalls – From First Alert to Effectiveness Checks. U.S. Food and Drug Administration [online]. Available: http://www.fda.gov/ForConsumers/ConsumerUpdates/ucm049070.htm [accessed Oct. 22, 2010].

FDA (U.S. Food and Drug Administration). 2010b. Manufacturer and User Facility Device Experience Database (MAUDE). U.S. Food and Drug Administration [online]. Available: http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/PostmarketRequirements/ReportingAdverseEvents/ucm127891.htm [accessed Nov. 8, 2010].

FDA (U.S. Food and Drug Administration). 2011a. FDA's Sentinel Initiative. U.S. Department of Health and Human Services, Food and Drug Administration [online]. Available: http://www.fda.gov/Safety/FDAsSentinelInitiative/ucm2007250.htm [accessed Apr. 18, 2011].

FDA (U.S. Food and Drug Administration). 2011b. FDA-TRACK CDRH Office of Surveillance and Biometrics (OSB) Dashboard. U.S. Department of Health and Human Services, Food and Drug Administration [online]. Available: http://www.fda.gov/AboutFDA/Transparency/track/ucm203271.htm [accessed Apr. 18, 2011].

Gill, G.S., D. Mills, and A.B. Joshi. 2003. Mortality following primary total knee arthroplasty. J Bone Joint Surg. Am. 85-A(3):432-435.

INTERMACS (Interagency Registry for Mechanically Assisted Circulatory Support). 2010. Quarterly Statistics Report: Implant Dates March 1, 2006 – June 30, 2010. Interagency Registry for Mechanically Assisted Circulatory Support [online]. Available: http://www.uab.edu/ctsresearch/intermacs/Document%20Library/INTERMACS%20Federal%20Partners%20Quarterly%20Report%20June%202010%20website.pdf [accessed Nov. 4, 2010].

Katz, B.P., D.A. Freund, D.A. Heck, R.S. Dittus, J.E. Paul, J. Wright, P. Coyte, E. Holleman, and G. Hawker. 1996. Demographic variation in the rate of knee replacement: A multi-year analysis. Health Serv. Res. 31(2):125-140.

Keefe Bartels, LLC. 2010. Smith & Nephew Oxinium Knee Replacement. Keefe Bartels, LLC [online]. Available: http://www.keefebartels.com/CM/HotTopicsandAlerts/Smith-Nephew.asp [accessed Nov. 8, 2010].

Kirklin, J.K., D.C. Naftel, R.L. Kormos, L.W. Stevenson, F.D. Pagani, M.A. Miller, K.L. Ulisney, J.T. Baldwin, and J.B. Young. 2010. Second INTERMACS annual report: More than 1,000 primary left ventricular assist device implants. J. Heart Lung Transplant. 29(1):1-10.

Mahomed, N.N., J. Barrett, J.N. Katz, J.A. Baron, J. Wright, and E. Losina. 2005. Epidemiology of total knee replacement in the United States, Medicare population. J Bone Joint Surg Am. 87(6):1222-1228.

Moss, A.J., S. Hamburger, R.M. Moore, Jr., L.L. Jeng, and L.J. Howie. 1991. Use of Selected Medical Device Implants in the United States, 1988. Advance Data No. 191. DHHS (PHS) 91-1250. U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control, National Center for Health Statistics, Hyattsville, MD.

NCHS (National Center for Health Statistics). 2010. Deaths and Mortality. Centers for Disease Control, National Center for Health Statistics, Hyattsville, MD [online]. Available: http://www.cdc.gov/nchs/fastats/deaths.htm [accessed Nov. 8, 2010].

Palmer, S.H., and M.J. Cross. 2010. Total Knee Arthroplasty. eMedicine Orthopedic Surgery-Knee [online]. Available: http://emedicine.medscape.com/article/1250275-overview [accessed Oct. 26, 2010].

Parvizi, J., T.A. Sullivan, R.T. Trousdale, and D.G. Lewallen. 2001. Thirty-day mortality after total knee arthroplasty. J. Bone Joint Surg. Am. 83-A (8):1157-1161.

Paxton, E.W., R.S. Namba, G.B. Maletis, M. Khatod, E.J. Yue, M. Davies, R.B. Low Jr., R.W. Wyatt, M.C. Inacio, and T.T. Funahashi. 2010. A prospective study of 80,000 total joint and 5000 anterior cruciate ligament reconstruction procedures in a community-based registry in the United States. J. Bone Joint Surg. Am. Suppl 2:117-132.

Slaughter, M.S., J. Rogers, C.A. Milano, S.D. Russell, J.V. Conte, D. Feldman, B. Sun, A.J. Tatooles, R.M. Delgado III, J.W. Long, T.C. Wozniak, W. Ghumman, D.J. Farrar, and O.H. Frazier. 2009. Advanced heart failure treated with continuous-flow left ventricular assist device. N. Engl. J. Med. 361(23)2241-2251.

Taylor, H.D., D.A. Dennis, and H.S. Crane. 1997. Relationship between mortality rates and hospital patient volume for Medicare patients undergoing major orthopaedic surgery of the hip, knee, spine, and femur. J. Arthroplasty 12(3):235-242.

Terracciano, C.M., L.W. Miller, and M.H. Yacoub. 2010. Contemporary use of ventricular assist devices. Annu. Rev. Med. 61:255-270.

UAMS (University of Arkansas for Medical Sciences). 2010. Inpatient Mortality Rate for Knee Replacement. University of Arkansas for Medical Sciences [online]. Available: http://www.uamshealth.com/?id=1240&sid=1 [accessed Oct. 26, 2010].

6

Case Study of a Targeting Decision That Spans Food and Drug Administration Centers

This final case study focuses on the use of the proposed framework to set priorities for work that could affect more than one center when choices must be made to allocate agency resources to several pressing needs that arise simultaneously. It is based loosely on two decision scenarios provided by the Food and Drug Administration (FDA), one of which described the need to set priorities for laboratory resources and the second described the need to balance a flood of laboratory testing requests associated with the 2007 melamine-contamination threat with other demands on laboratory resources. As with the other case studies, the data used were gleaned from publicly available Web sites or publications or were provided by FDA. The committee did not conduct exhaustive literature searches or reviews, and all information is illustrative. The case study simply provides an illustration of how the committee's framework might be used for a targeting decision that spans FDA centers.

FRAMING THE ISSUE: MELAMINE CONTAMINATION

This case study considers a hypothetical situation in which several products are considered to be at risk for melamine contamination, and a series of decisions must be made about how to set testing priorities to understand the extent of contamination. The case study is built on events that occurred in 2007 and 2008. Some of the key issues associated with potential melamine contamination and the chronology of events that took place are described below.

Melamine and Melamine Contamination in Human Food and Animal Feed

Melamine is a common stable chemical used to produce a variety of materials, including resins, laminates, glues, adhesives, coatings, and flame retardants (WHO 2008). It is a residual byproduct of metabolism of the pesticide cyromazine in plant and animal tissues (Lim et al. 1990; FAO/WHO 2010). Cyromazine is not approved for use in the United States but has approval in other countries that might export plant or animal products. Melamine contains relatively high nitrogen content (66% by mass), and this property has been exploited by some to raise nitrogen concentrations in products tested with conventional nitrogen-detection methods. Specifically, standardized protein-test methods rely on nitrogen concentration as a proxy for protein content, so the high nitrogen content of melamine can be used to increase the tested protein content of a food artificially.

Melamine and its analogues have been found to have low oral toxicity in laboratory animals. Short-term exposures to melamine require high doses (LD_{50}, about 3 g/kg in rats) to cause acute toxic effects (WHO 2009). The toxic effects appear to be restricted primarily to the kidney and bladder.

Despite its low toxicity in laboratory animals, melamine has been associated with several outbreaks of toxicity in humans and companion animals. In animals, the consumption of feed that contains melamine and cyanuric acid has been shown to cause nephrolithiasis (kidney stones) and renal failure at much lower doses than would occur with either chemical alone (Puschner et al. 2007).

China is recognized as the leading manufacturer of industrial melamine in the world and uses it in a wide array of products. One of the byproducts of industrial melamine production is a less concentrated version known as melamine scrap. Melamine scrap is reported to contain impurities, including cyanuric acid, which may increase its potential toxicity in humans and animals (Bradley 2008a,b). Recently, it has been reported that melamine scrap has been used in abundant quantities for years as a "protein powder" to boost the nitrogen values of livestock feedstuffs produced in China (Reuters 2008). Furthermore, similar protein powder concoctions have allegedly been added directly to liquid milk to ensure that the milk passes protein-quality standards (Ma 2008).

Timeline of Events

Melamine made headlines in the United States in 2007 when numerous pets presented with illnesses related to renal malfunction or failure and sudden death. Initially, the widespread occurrences were sufficiently disparate in location, food sources, and presentation as not to manifest a singular clinical picture of food toxicity. However, as the number of affected dogs and cats rose above an estimated 1,000, veterinary diagnostic laboratories and FDA became involved in investigating the root cause of the clinical presentations. As the search for causation narrowed, melamine and cyanuric acid were identified in pet-food samples received from owners of the animals that had toxicosis. Multiple pet-food manufacturers determined that pet food that they produced and marketed was adulterated with melamine and cyanuric acid, and they instituted recalls. The contamination was judged to be the cause of the clinical manifestations in affected dogs and cats and was consistent with an earlier outbreak in Asian dogs and cats reported in the literature (Brown et al. 2007). Ultimately, over 150 brands of dog and cat foods were associated with the contamination.

Further investigation of the event found that wheat gluten and rice-protein concentrate that originated in China were the sources of pet-food adulterants. Food-microscopy analyses indicated that melamine crystals had been used to spike the protein concentrations to acceptable commercial standards. As a result of publicity about the investigation, a separate source of melamine was found in selected domestically produced animal and aquatic feeds (Itchmo 2007; Tembec BTLSR Inc. 2007). Two companies were selling a pellet binder in aquatic feeds that contained low melamine concentrations, both domestically and internationally, and products were withdrawn from U.S. markets on discovery of the contamination. That event demonstrated that potential contamination was not limited to products directly from China. Therefore, investigations of other feed ingredients, both imported and domestic, were appropriate to ensure the safety of pet foods and food-animal feeds that could be manufactured with such ingredients.

In September 2008, news from China indicated foodborne toxicosis in young children caused by melamine contamination of infant formula. On December 1, 2008, the Chinese Ministry of Health reported that over 290,000 children were being treated for clinical signs of toxicosis; 51,900 children were hospitalized, and six deaths were confirmed (Gossner et al. 2009). An investigation by the Centre for Food Safety in Hong Kong indicated that 99% of children who exhibited clinical signs were less than 3 years old (CFS 2008). Investigations indicated that likely sources of contamination were dried dairy products made from milk produced in China. It was determined that milk was being diluted with water and that melamine was being added to increase the nitrogen (protein) concentration to hide this diversion. Because the practice was reported to be relatively widespread in rural China (an estimated 20% of dairies in the country were involved), large quantities of milk and milk products were potentially contaminated. Baby formula was the first product to be identified because of the toxic effects observed in children exposed to contaminated brands of formula. However, a wide array of other foods—including cookies,

biscuits, confections, milk-flavored instant coffee and tea, and other products—that contain powdered milk as an ingredient were suspected of melamine contamination.

DECISION CONTEXT FOR THE CASE STUDY

For this case study, the committee assumed that an analytic laboratory receives many demands from field investigators for melamine testing in two products: infant formula and animal-feed ingredients. Extensive testing of any one of the product types, using the testing methods available, would overwhelm available laboratory resources. There is strong circumstantial evidence that one or more products may be at risk for contamination, but the sources and extent of contamination are unknown at the outset. Before other FDA resources are redirected or outside laboratories are engaged, some sense of the magnitude of the problem must be ascertained. This is an example of a targeting decision: FDA must decide where, among competing demands, it should allocate resources when all priorities cannot be addressed simultaneously. Follow-up decisions could include both mitigation-selection decisions, in which FDA decides which (if any) actions the agency will take to reduce or mitigate the identified risks, and, in the longer term, strategic-investment decisions, in which FDA decides whether to invest resources in improved testing methods, increased surveillance, or similar efforts devoted to understanding the identified risks.

The case study examines the first type of decision (a targeting decision) in two steps. The first step is to characterize and compare the public-health consequences associated with potential contamination of the two products. That example is fully developed through the evaluation and completion of the attribute table. The second step is to consider how laboratory resources should be allocated given the understanding of risk magnitudes from the comparison. The approach that one could take to addressing the second step is described but is not fully developed.

CHARACTERIZING THE PUBLIC-HEALTH CONSEQUENCES

For this case study, the committee attempted to rely on information that would have been available to FDA laboratory personnel at the time of the historical event in 2007-2008. The products discussed in this case study—infant formula and animal-feed ingredients—are credible sources of melamine in 2008; at that time, the potential for melamine in liquid milk was not known. In characterizing the public-health consequences, the committee relied on general, publicly available information about the population size and food consumption patterns, as described below. The melamine-specific information that was needed to develop the case study included estimates of the probability and concentration of melamine in each food type and the human dose-response relationship. Those factors were highly uncertain, and the committee struggled to find descriptions of the data that would have been available to FDA in 2008. In developing the detailed estimates described below, the committee consulted two primary sources. First, melamine concentrations found in infant formula in China in 2008, as reported in Gossner et al. (2009), were used as a basis of estimates of the potential concentrations in formula. Second, the study *Interim Safety and Risk Assessment of Melamine and Its Analogues in Food for Humans* (FDA 2008) provided the basis of most of the assumptions about the human risk associated with contaminated animal feeds and the information to hypothesize a dose-response relationship.

The lack of data and the complexity of the situation being evaluated required the committee to make informed judgments, assumptions, and a variety of calculations to characterize the public-health consequences of potential contamination of the two products. Those judgments, assumptions, and calculations are described below. Table 6-1 summarizes the risk characterization in terms of the attributes.

TABLE 6-1 Risk Attributes for Targeting Decision that Spans Food and Drug Administration Centers

Attribute	Metric	Product type	
		Infant Formula	Hog-Feed Ingredients
Exposed population	Number exposed to potential melamine contamination	160,000 (60,000-320,000)	500,000 (240,000-1,700,000)
	Populations of concern	Entire exposed population consists of children under the age of 1 year	No populations of special concern identified
Mortality	Number of deaths per year	$6x10^{-5}$ (0-0.04)	0
Morbidity	Number experiencing severe adverse health effects per year	1 (0-630)	0
	Number experiencing less severe adverse health effects per year	4.5 (0-2,900)	0
	Number per year experiencing adverse health effects that affect only quality of life	5.5 (0-3,500)	0
Personal controllability	Degree to which a person can eliminate or reduce his or her own risks through voluntary actions	Before any determination that melamine might be present in infant formula, parents have no ability to control risks to their infants; once aware of the potential problems, 90-95% of parents could avoid potentially contaminated formula.	Before any determination that melamine might be present in pork products, consumers have no ability to control those risks; once aware of the potential problem, 80-90% of consumers could avoid potentially contaminated food products.
Ability to detect adverse health effects	Ability to detect unexpected population-level adverse effects	Adverse health effect rates would have to be 10-100 times higher than expected to be detected and correctly attributed.	Not detectable
Ability to mitigate adverse health effects	Probability that an informed institution will be able to reduce or mitigate any adverse health effects associated with the product being evaluated if such a problem is known to exist	90-100%	10-50% of effects associated with contaminated pork products could be mitigated before the contamination source is identified; 90-100% of future problems could be mitigated once the contamination source is known.

Infant Formula

Exposed Population

About 4 million children are born in the United States each year, and most of them consume at least some formula during the first year of life. Those younger than 6 months old are at greatest risk from contaminated formula because their primary dietary sources are exclusively baby formula or baby formula and breast milk.[1] After 6 months, most children begin to reduce formula consumption as their diet shifts to cereals and other solid foods; after 1 year, most children have transitioned away from formula. Table 6-2 shows the number of children in an annual birth cohort that fall within various age groups, the percentage of children in each age group who consume infant formula at least once a day, and the resulting number of children annually who consume infant formula at least once a day.

However, the committee defined the exposed population for this case study as consisting of children consuming infant formula that is potentially contaminated with melamine, not simply all children consuming baby formula. U.S. baby formula is produced by five prominent manufacturers (over 90% of the market) that, when surveyed by FDA, reported exclusive use of domestic milk products in their formulations. None reported use of imported milk-protein sources in their formulations. Imported baby formula and possibly baby formula from nonsurveyed domestic producers would have been suspect. On the basis of that information, the committee estimated that 2-10% of the total U.S. infant formula supply would have had the potential for some level of melamine contamination; the best estimate was 5% of the supply.

Assuming that the market and consumption patterns are such that an infant's formula source does not vary substantially over time (that is, there is high brand loyalty), the best estimate of the number of infants exposed to potentially contaminated formula is about 161,000 (5% of 3.23 million) with a range of about 62,000-320,000.

TABLE 6-2 Formula Consumption by Age

Age Group	Number of Children in Given Age Group	Percentage Who Consume Infant Formula at Least Once a Day	Number of Children Who Consume Infant Formula at Least Once a Day
Less than 4 months	1,333,000	83	1,110,000
4-6 months	1,000,000	74	740,000
7-8 months	667,000	82	547,000
9-11 months	1,000,000	75	750,000
12 months	333,000	5-50 (25 best estimate)	16,700-167,000
Total (birth to 12 months)			3.23 million (best estimate) 3.16 million to 3.31 million

Note: Data from Briefel et al. (2004) cover only ages 4-11 months. Percentage of children under 4 months old who consume formula is based on percentage reporting breast milk only at age of 4 months. Formula consumptions by children 12 months old is assumed to scale back substantially as they transition to cow's milk.

[1]The National Immunization Survey (CDC 2010) found that 13.6% of infants born in 2006 were exclusively breastfed through the age 6 months; these infants would not be at risk from potential melamine contamination in baby formula.

Mortality and Morbidity

Although melamine is not regarded as highly toxic, continuous exposure to melamine-contaminated formula could result in nephrolithiasis in a portion of the exposed population. On the basis of the Chinese experience in 2008, the mortality rate is likely to be low. However, any estimate is highly uncertain because of the lack of historical data and a poor understanding of the dose-response relationship, the effect that other contaminants (such as cyanuric acid) might have on overall toxicity,[2] consumption patterns, and other factors, such as cultural practices, that might affect overall toxicity. Rough estimates of the number of deaths that could result from melamine contamination in infant formula can be calculated as follows:

$$\text{Number of deaths} = \int_{d=0}^{\max \text{dose}} \Pr[death|d] \times N_d, \quad (6\text{-}1)$$

where d represents dose, and the integration is over the entire range of potential doses; Pr[*death*|*d*] represents the dose-response relationship as the probability that a person who receives dose d will die; and N_d is the number of infants who receive that dose. The number who would experience each type of health effect (severe adverse effects, less severe effects, or adverse quality-of-life effects) can be estimated similarly by substituting the probability of that effect as a function of dose for Pr[*death*|*d*] in Equation 6-1.

The critical factors in the calculation are the number of children who receive various doses and the dose-response relationship. Each factor is highly uncertain, but the uncertainty can be estimated or modeled. The discussion below illustrates how a relatively quick analysis of the factors might be conducted; more rigorous analyses could be conducted with more time, effort, and data.

Dose Estimates

Estimating the number of children who receive various doses of melamine through contaminated formula requires estimates of each of the following:

- Total number of infants who consume potentially contaminated infant formula (N_d as shown above).
- Level of contamination in the formula.
- Daily intake of formula relative to body weight.

At the time of this hypothetical study, little was known about the extent of melamine contamination in U.S. infant formula, that is, the amount of formula contaminated or the melamine concentration in any contaminated formula. (For this case study, it is assumed that all concentrations shown in parts per million are for the formula powder that is added to water, not the liquid that is consumed. Once mixed with water, formula constitutes about 10% of the liquid.) The World Health Organization (WHO 2009) reported concentrations up to several thousand parts per million in some of the Chinese formulas in 2008. Gossner et al. (2009) reported data that showed that the melamine-contaminated formula in 18 of 22 cases had concentrations less than 100 ppm. For the purpose of this study, the concentrations in contaminated dry formula in the United States are uncertain and assumed to be 0-200 ppm with 10 ppm as a most likely value. The committee notes that this analysis could be easily re-run for different assumptions as to the range of possible values.

[2]Studies released after the timeframe for this hypothetical case study clearly identified the importance of the presence of cyanuric acid with melamine for manifestation of the toxic effects observed. However, the committee assumed that the strength of that relationship was still uncertain at the time of this case study.

Formula consumption varies with age, body weight, and infant eating habits, as shown in the first four columns of Table 6-3 (Family Education 2010). It is important to note that the amount of formula consumed begins to decrease after the age of 3-6 months as other foods are added to children's diets. Calculating the melamine dose to the exposed population requires combining the estimated concentration of melamine in the formula and the amount of formula consumed relative to body weight. Table 6-3 shows an example calculation of the dose in milligrams per kilogram of body weight per day based on an assumed concentration of 100 ppm in dry formula (diluted to 10 ppm in liquid formula) and the body weight and formula consumption estimates shown.

Table 6-3 shows, for example, that a 1-week-old consuming formula contaminated at 100 ppm in the dry formula will most likely ingest melamine at 1.5 mg/kg per day, but it could be as high as 2.6 or as low as 0.5 mg/kg per day depending on eating habits and infant weight. After the time that this hypothetical study would have been conducted, WHO set the tolerable daily intake (TDI) at 0.2 mg/kg per day (WHO 2009). The committee assumed that the information used to support that determination would have been available at the time of the study.

Table 6-3 shows the range of dose for a single concentration of melamine in dry formula (100 ppm) for illustrative purposes; the actual concentration is uncertain. Including uncertainty about the concentration increases the uncertainty of the estimated dose received by the exposed group. For example, for a single exposed child 1-3 months old, assumed to be of average weight (6 kg) and consuming an average amount of formula per day (31 oz), a concentration of 100 ppm in dry formula leads to a dose of 1.46 mg/kg per day, as shown in Table 6-3. However, if uncertainty in the concentration in formula is included in the calculation (that is, a distribution with 10 ppm as the best estimate and a range of 0-200 ppm), the estimated dose for that child is now a distribution with a median of 0.15 mg/kg per day and a range of 0-2.9 mg/kg per day.

Repeating that calculation for all the uncertainties described above is easily done with a simple Monte Carlo simulation model. Such a model[3] was developed to generate an estimate of the daily intake of melamine by an exposed infant. Figures 6-1 and 6-2 illustrate the results of the calculation. Figure 6-1 shows the median and 5th and 95th percentiles of the estimated intake by infants in each of six age groups, incorporating uncertainties in melamine concentration in formula and amount of formula consumed in each age group. The uncertainty is greatest for the youngest children, and both the median intake and the uncertainty in the intake decrease for children over 3 months old as their formula consumption decreases and body weight increases. For comparative purposes, the figure also shows the WHO TDI of 0.2 mg/kg per day. The median intake estimated for all age ranges is below the TDI, but the 95th percentile estimate is significantly higher. Figure 6-2 provides the full distribution of intake, displayed as an inverse cumulative distribution function (CDF). The inverse CDFs for various infant ages show the probability of exceeding different doses. For example, looking at a dose of 1 mg/kg per day on the x axis and tracing up to the line for children from birth to the age of 3 months, one can see that there is a 0.2 probability (20% chance) that a child in that age range consuming contaminated formula will have a dose that exceeds 1 mg/kg per day. The probability that a child 9-12 months old (the lowest line on the figure) would exceed 1 mg/kg per day is lower, about 0.05.

The final step in estimating the number of infants receiving each dose is to combine the number of exposed children in each age group (which is also uncertain) with the distribution of dose by age shown above and then aggregate across the age groups. Figure 6-3 shows the distribution of the total number of infants exceeding specified doses measured in milligrams per kilogram per day. For example, as shown in the figure, about 70,000 infants every day consume melamine at 0.1 mg/kg or more; the number could be as low as 0 or as high as 290,000.

[3]For this case study, a simulation was run with "@Risk," an Excel Add-In by Palisade Corporation. Uncertainty in concentration in formula was modeled with a gamma distribution with parameters set to match the 5th, 50th, and 95th percentiles of concentration described in this report; all other uncertainties were represented by using triangular distributions also defined so that the 5th, 50th, and 95th percentiles match the estimates described in this report.

TABLE 6-3 Estimated Dose for Infants Consuming Formula with Melamine at 100 ppm in Dry Formula

	Formula per day, oz			Melamine consumed per day, mg		Daily dose of melamine, mg/kg		
Age	Min	Max	Average infant weight, kg	Min formula consump	Max formula consump	Min formula consump	Max formula consump	Median
1 week	6	30	3.3	1.70	8.51	0.52	2.58	1.55
1 week to 1 month	14	32	5	3.97	9.07	0.79	1.81	1.30
1-3 months	20	42	6	5.67	11.91	0.95	1.98	1.46
3-6 months	24	35	7.8	6.80	9.92	0.87	1.27	1.07
6-9 months	21	32	9.2	5.95	9.07	0.65	0.99	0.82
9-12 months	21	24	10.2	5.95	6.80	0.58	0.67	0.63

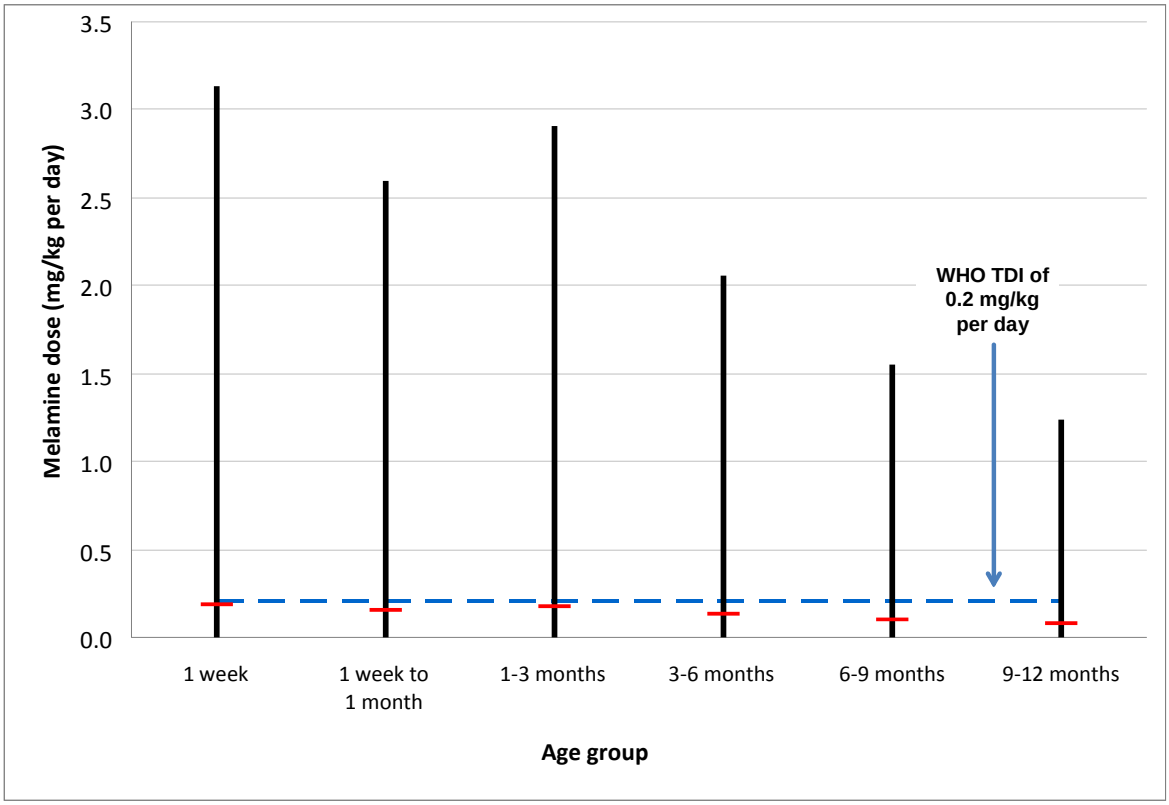

FIGURE 6-1 Median (red dash) and range (black vertical line) of estimated melamine dose for an exposed infant of various ages given uncertainties in melamine concentration in formula, infant weight, and formula consumption. Blue dashed line represents tolerable daily intake (TDI) from the World Health Organization (WHO).

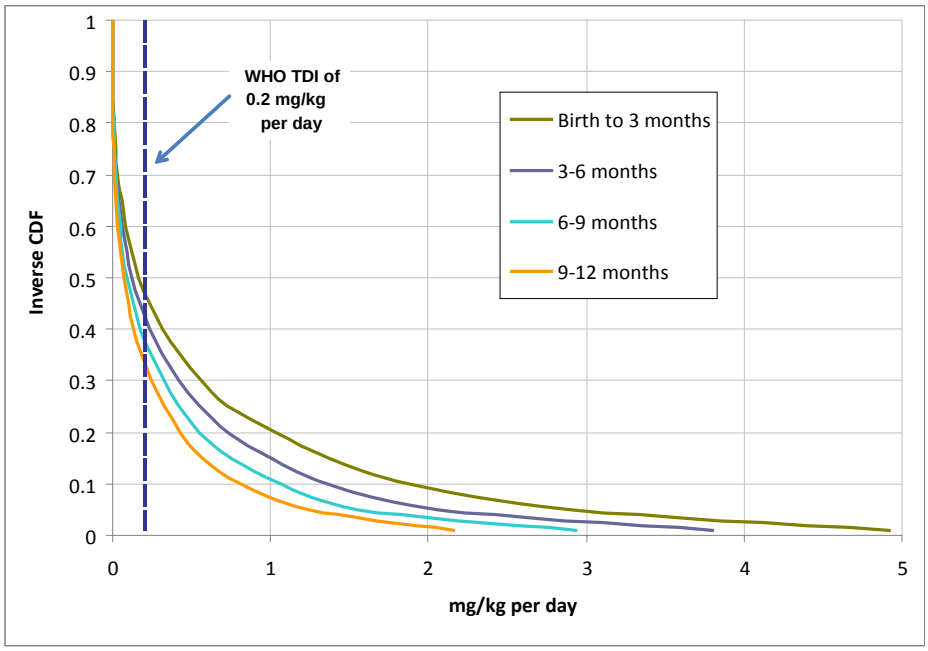

FIGURE 6-2 Distribution of estimated melamine dose for an exposed infant of various ages given uncertainties in melamine concentration in formula, infant weight, and formula consumption. CDF = cumulative distribution function. Blue dashed line represents tolerable daily intake (TDI) from the World Health Organization (WHO).

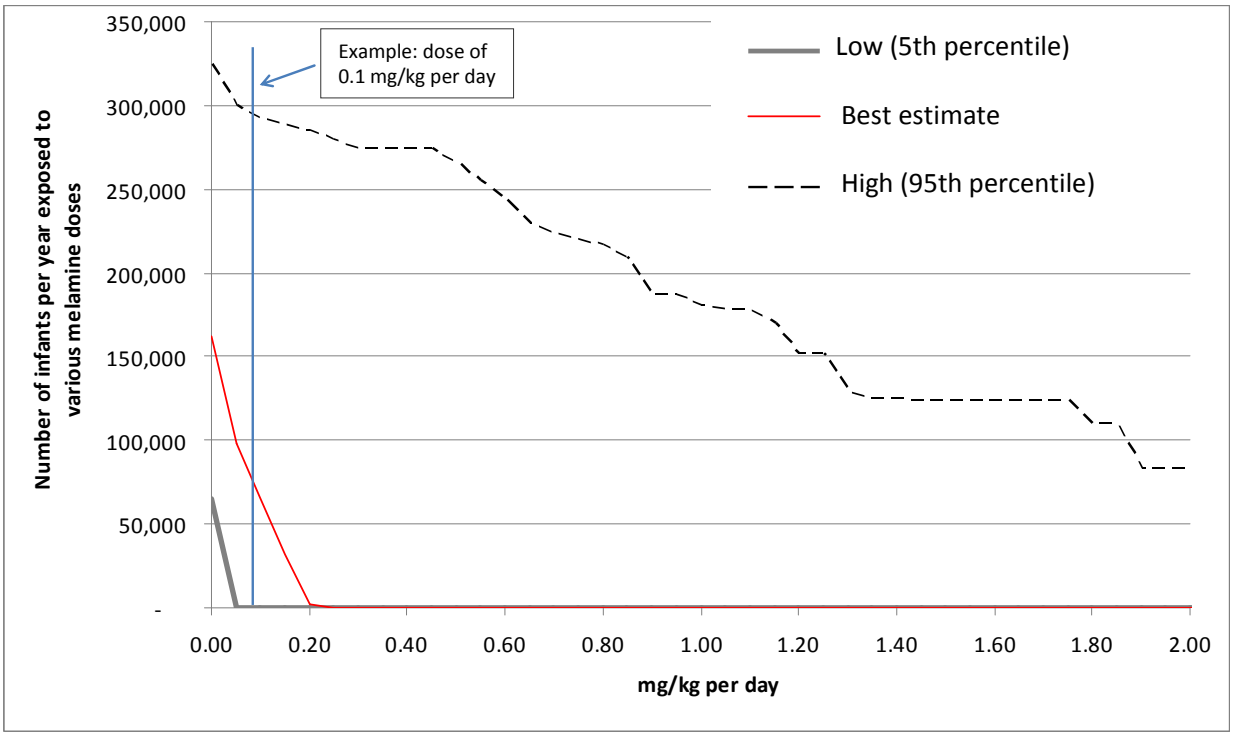

FIGURE 6-3 Distribution of the number of infants exposed to various melamine doses. Red line shows the median estimate of the number of infants who receive at least the indicated dose. Dashed and gray lines show the high and low estimates, respectively.

Dose-Response Relationship

At the time this study would have been conducted, little was known about the dose-response relationship of the effects observed in infants and melamine consumption. In China, where melamine concentrations ranged from less than 1 ppm to over 1,000 ppm, Gossner et al. (2009) stated that there were 294,000 cases of renal illness, including 51,900 that required hospitalization, and at least six deaths. The number of infants that were exposed to the high concentrations is not known, so precise risk calculations are impossible, but the ratios of deaths to hospitalizations to less severe renal disease is useful. Lack of historical and scientific data on the long-term effects of infant morbidity and on lesion severity makes characterizing the health effects of the renal illnesses equally difficult and highly uncertain. Depending on the melamine dose, the frequency of exposure, and the percentage of formula affected, a substantial part of the cohort might be involved. However, if high concentrations were present in formula in the United States, sick infants would soon start appearing in doctors' offices and emergency rooms, and there would be a quick response, so repeated exposure to large amounts is not expected; this is described in following sections that discuss mitigation opportunities.

A hypothesized dose-response relationship between melamine dose and the percentage who would experience adverse health effects at that dose is shown in Figure 6-4. The actual values are highly uncertain, and the hypothesized relationship is intended only to provide a starting point for further analyses. The estimated distribution of the types of effects (that is, the ratios of deaths to severe adverse effects to less severe adverse effects) is based loosely on the data from China, assuming that the ratios are essentially constant over exposure and that for every infant who experiences clinical signs of renal disease another suffers from discomfort or some other effect that diminishes quality of life. The overall estimate of the rate of any adverse effects is consistent with an assumption that if exposure is 10 times the WHO TDI of 0.2 mg/kg per day, less than 5% of the population would experience any adverse effects. The dose-response relationship shown here—given the assumption that response can be extrapolated linearly as a function of dose—produces the same magnitude of adverse effects as occurred in China if 4 million infants consumed formula adulterated with melamine at 200 ppm in liquid form, a concentration substantially higher than the concentrations hypothesized in this case study.

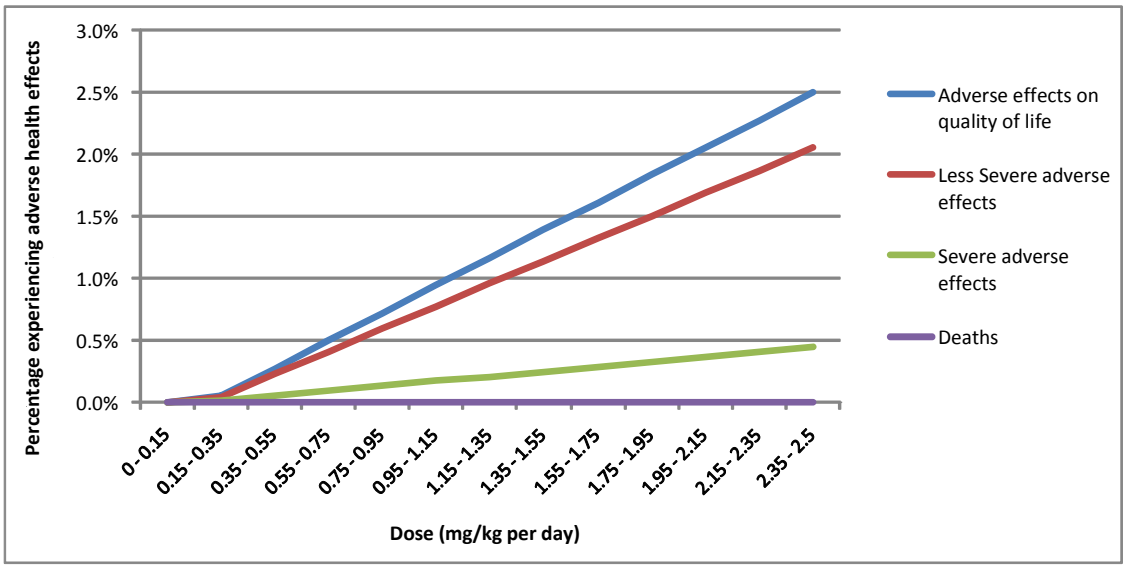

FIGURE 6-4 Dose-response relationship for health effects of melamine consumption used for illustrative calculations.

Number of Deaths and Other Adverse Health Effects

The distribution of health outcomes can be estimated by using the logic of Equation 6-1 and the estimated dose distribution described above with the dose-response relationship. Table 6-4 shows the results of that calculation. The best estimates are that the total number of adverse health effects of any sort would be around 10, and that deaths would be extremely unlikely. At the high end, it is possible that thousands of cases of renal disease requiring medical attention would occur; deaths would still be very rare.

Personal Controllablity

To be able to control the risks to their children from potentially contaminated formula, parents must be aware that risks exist and must have the ability to exercise options to avoid or reduce the risks. In this case study, the main issue is whether parents are aware that a small portion of the total supply of infant formula is potentially contaminated. There is no way to detect contamination on the basis of a physical examination of the product itself, so such awareness would have to come from a determination and an announcement of potential contamination from a public-health agency. At the time this study would have been conducted, it was not known whether such contamination existed, so parents of the potentially exposed children would have had virtually no ability to control the risks. Once aware of the potential health consequences, parents could take steps to avoid potentially contaminated formula relatively easily. If the problem were isolated to a small number of identified formula manufacturers, the availability of formula from other manufacturers would allow virtually all parents to select uncontaminated formula.

Ability to Detect Adverse Health Effects

The probability that a single case of nephrolithiasis caused by melamine-contaminated infant formula would be detected depends on the severity of the illness. According to the American Society of Pediatric Nephrology, signs and symptoms of melamine poisoning in infants include the following:

- Unexplained crying in infants, especially when urinating.
- Vomiting.
- Unexplained fever due to urinary tract infection caused by urinary obstruction.
- Blood in the urine.
- Stones discharged while urinating.
- Signs of renal infection.
- High blood pressure.
- Elicitable flank pain.

TABLE 6-4 Estimates of Mortality and Morbidity from Melamine Contamination in Infant Formula for the Hypothesized Dose-Response Relationship Incorporating Uncertainty in Exposure and Dose

Type of Effect	Low	Median	High
Affecting quality of life only	0	5.5	3,500
Less severe adverse effects	0	4.5	2,900
Severe adverse effects	0	1.0	630
Death	0	6×10^{-5}	0.04

Detecting population-level effects resulting from melamine contamination of infant formula at the hypothesized level and correctly attributing the cause of the effects would be difficult. The presence of kidney stones in infants would not necessarily trigger concern about melamine contamination, but if multiple children developed kidney stones with no other risk factors, the problem would probably be detected. However, the ability of institutions to detect a problem, assuming that one exists, would probably be somewhat hindered because the contaminated formula would likely be geographically dispersed and the vast majority of adverse health outcomes would be minor.

In the case of the melamine-contaminated pet food in 2007, it was estimated that by the time the problem was identified, about 100 pets had died and about 500 had suffered kidney failure (AP 2007); some estimates of the number of deaths exceeded 4,000 (Weise and Schmit 2007). The committee assumed that unexpected cases of kidney stones in infants would be recognized as a problem much sooner than adverse health effects in pets. Thus, the committee estimated that adverse effects occurring 10-100 times more frequently than expected would be detected (for example, 50-500 children with kidney stones serious enough to require medical attention would be a sufficiently strong signal to health authorities that there was a systematic problem).

Ability to Mitigate Adverse Effects

In this framework, the ability to mitigate refers to the ability to manage, reduce, or otherwise control any adverse health effects of the products being evaluated, assuming that such effects occur and are detected. The manufacture of infant formula is well understood and controlled, although its distribution and consumption are less so. As discussed in Chapter 4, the effectiveness of a recall depends on several factors. However, the committee assumes that the recall of infant formula given the population potentially affected would be much more effective than, for example, the recall of canned foods. There would most likely be a highly visible media campaign to alert parents. Furthermore, because the potentially exposed population receives frequent medical check-ups, pediatricians would certainly alert parents to the potential dangers. Therefore, the committee assumes that recall procedures would be quick and effective and that the prompt notification of wholesale and retail outlets and of the general public would lead to a high probability of mitigation (90-100%).

Animal-Feed Ingredients

The second product of interest for this case study is animal-feed ingredients. There is potential for melamine to enter the human food supply through contaminated animal feeds. That is, if animals raised for human consumption eat feed that is contaminated with melamine, melamine may accumulate in the animal's flesh and ultimately be eaten by humans. The potential exposure pathway for animals is a feed ingredient that is adulterated with melamine to disguise protein deficiencies; this was the case with some pet foods. Hogs are sometimes fed excess (scrap) pet food as an inexpensive addition to their diets. Because melamine contamination of pet food was known to have occurred, the exposure pathway considered most likely and evaluated here is from contaminated pet foods being fed to hogs raised for human consumption.

When melamine contamination in pork was first suspected, each step in the pathway from contaminated animal feed to human consumption involved great speculation and uncertainty. To clarify the situation, a series of studies and laboratory tests were conduced that resulted in a 2007 report, *Interim Melamine and Analogues Safety/Risk Assessment*, that greatly reduced the concerns associated with this pathway (FDA 2007). The discussion presented here is based on what was known and believed at the time that report was written.

Exposed Population

The exposed population is defined as the number of people in the United States who could consume pork that is contaminated with detectable melamine. That calculation requires estimating the size of the population that consumes pork and the percentage of the total pork supply that could be contaminated with melamine.

The U.S. per capita pork consumption has been steady, averaging 1 lb/week, and about 80% of the U.S. population regularly consume some form of pork products (FDA 2007). Only a very small portion of the U.S. hog supply is fed excess pet food, the only hypothesized route to melamine contamination. In 2010, the U.S. hog and pig inventory was 64 million head (USDA 2010). Over the last 2 decades, the number of hog farms has been sharply reduced—from over 650,000 farms in 1980 (NHF 2005) to 74,000 in 2009 (USDA 2009)—as the size of the farms has dramatically increased. Table 6-5 shows the concentration of production on large farms (over 40% of hogs are produced on 0.1% of the farms). The large-scale operations are highly efficient and tightly controlled operations and are unlikely to supplement their animal feed with excess (scrap) pet food. For this analysis, the problem focuses on the numerous smaller farms that make up a few percent of the market where operations are less regimented and some low-cost excess pet food could be added to the animals' diet.

On the basis of expert judgment, the committee estimated that less than 1% of the pork that makes it to market would come from hogs that were fed pet food contaminated with melamine (0.001% of U.S. hogs is the low estimate, 0.2% is the median, and 0.7% is the high estimate). If 80% of the U.S. population of 310 million is assumed to eat pork, and 0.2% of the market consists of hogs that may have been exposed to melamine through consumption of excess pet food, hundreds of thousands of people have the potential to be exposed to melamine-contaminated pork (240,000-1,730,000, with a best estimate of 496,000).

Mortality and Morbidity

Estimating the potential extent of the problem in terms of deaths and other adverse health effects that might result from contaminated animal feeds is a multistep process. It follows the same logic used to estimate the potential health effects of melamine in infant formula described above. Estimating the mortality and morbidity requires estimates of the amount of melamine that would be consumed by members of the exposed population and the long-term health effects of melamine in humans (the dose-response relationship). However, several extra steps are also necessary. For example, to estimate the amount of melamine consumed by one person who is assumed to consume pork that has been fed melamine-contaminated scrap pet food requires the following estimates:

- Amount of melamine present in the animal feed.
  - Melamine concentration in scrap pet food.
  - Fraction of animal diet that is scrap pet food.
- How melamine from pet food accumulates in animal flesh.
- Amount of contaminated pork consumed.

Each of those factors is uncertain; they are explored in order below.

The committee estimated the melamine concentration in the diet of a single hog by multiplying the melamine concentration in the pet food consumed by the fraction of the diet that was assumed to be pet food. For this case study, the committee estimated that the melamine concentration in the contaminated pet food could range from 100 ppm to 15,000 ppm with a median of 1,000 ppm. The diet of hogs that receive any contaminated scrap pet food would vary considerably. The committee estimated that most would receive the contaminated pet-food supplement as a small portion of their total lifetime diet: 2% of diet is the low estimate, 35% is the median, and 100% is the high estimate. So, for example, a

hog consuming 35% of its diet as pet food that is contaminated at 1,000 ppm would have a net concentration of 350 ppm melamine in its diet. The committee combined the uncertainties in diet composition and melamine concentrations in the feed with uncertainty in the number of hogs consuming any contaminated pet food in a simple Monte Carlo simulation model to estimate a distribution of the number of hogs consuming feed with different net concentrations of melamine. All factors were assumed to be uncorrelated.

Figure 6-5 illustrates the resulting distribution of the number of hogs consuming a diet with various net concentrations of melamine. It shows the number of hogs consuming diets with melamine exceeding different concentrations. For example, looking at a concentration of about 5,000 ppm on the x axis and tracing up to the "best estimate" line, one can see that about 65,000 hogs will be consuming melamine at over 5,000 ppm in their diets. The 5th and 95th percentile lines indicate that the number of hogs can range from 20,000 to 150,000.

TABLE 6-5 Concentration of Hog Production in 2006

Number of Head Marketed	Number of Farms	Percent of Farms	Percent of Market
Under 1,000	48,434	86.1	1
1,000-2,999	4,025	7.1	5
3,000-4,999	1,150	2	3
5,000-9,999	1,100	1.9	6
10,000-49,999	1,450	2.6	21
50,000-499,999	164	0.3	21
500,000+	27	0.1	43
TOTAL	56,350	100	100

Source: Adapted from Pork Checkoff 2009, p. 87.

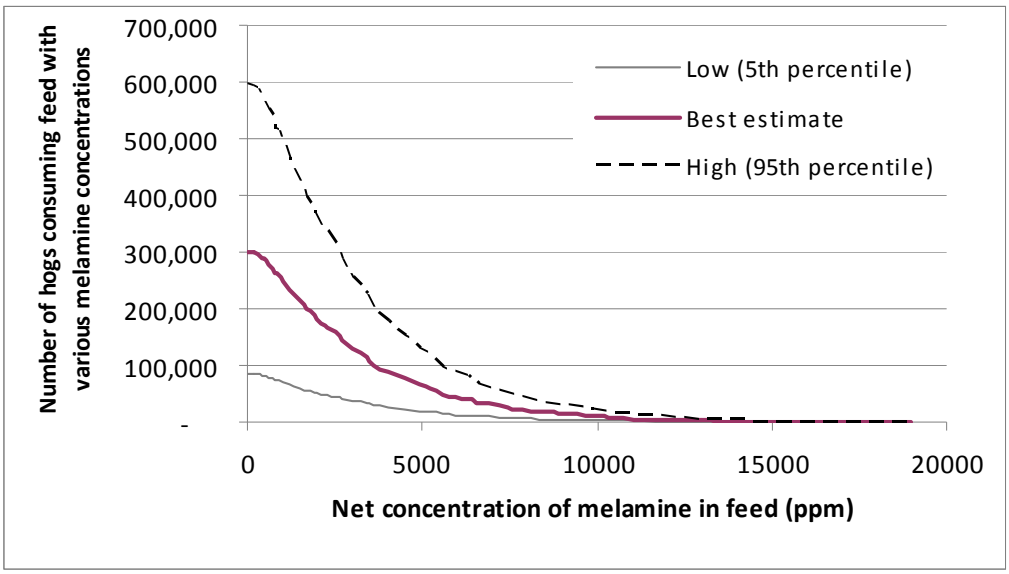

FIGURE 6-5 Distribution of hogs consuming feed with various melamine concentrations.

Given the melamine concentration in the feed, the next step is to estimate the melamine concentration in the animal tissue. On the basis of reports written after the contamination scare, the committee assumed that researchers believed that feed concentrations of less than 500 ppm would result in tissue concentrations less than the limit of detection of 50 ppb, or that tissue concentrations in the pork would be about 0.0001-0.001 net concentration in the food consumed by an animal. Combining the conservative factor of 0.001 with the estimated distribution of the number of hogs consuming foods with various concentrations shown in Figure 6-5 results in estimates of the numbers of hogs with various melamine concentrations in their tissues.

Figure 6-6 illustrates the results of the calculation. Hogs consuming melamine in their diet at a net concentration of less than 500 ppm are assumed not to have any detectable melamine in their tissues. Figure 6-6 groups the hogs that have detectable melamine into 1,000-ppb bins and shows the best estimate (the red marker) and the range of the number of hogs or pounds of pork product with those concentrations that are assumed to reach the market. For example, about 60,000 hogs (12 million pounds of pork) with melamine at 2,000-3,000 ppb would be expected, but the number could range from 20,000 to 120,000.

In the 2007 report on melamine risk, FDA estimated the consumption of pork and pork products that are contaminated with melamine at 100 ppb to deliver a dose of 0.04 µg/kg per day to the average person and 0.10 µg/kg per day to the 90th-percentile person. Extrapolating those values linearly with concentration to the distributions shown in Figure 6-6 yields the estimates shown in Table 6-6.

The last step in the multistep process requires estimating the human-health consequences of the doses described above. In its 2007 report on melamine risk, FDA specified a TDI of 0.63 mg/kg per day (630 µg/kg per day). In 2008, WHO established a lower TDI of 0.2 mg/kg per day (200 µg/kg per day). The TDI is the maximum amount that a human can be exposed to daily *over a lifetime* without incurring a measurable increase in health risk. The health effects associated with values above the TDI for melamine are kidney stones and renal disease. The highest dose estimated above (and shown in Table 6-6) is 15 µg/kg per day—1/13 the WHO TDI and 1/40 the FDA TDI. On the basis of that comparison, the facts that the TDI incorporates a safety factor of 100 and that a lifetime of contaminated pork consumption is not realistic given the history of the problem, the committee assumed that *no* adverse health effects would occur. In its preliminary analysis of risk, the committee did not include all the uncertainty, but even adding several layers of uncertainty to the hog or human population calculations would not push the risk to a level at which adverse health effects would be expected; it would only increase the size of the population exposed to low doses of melamine.

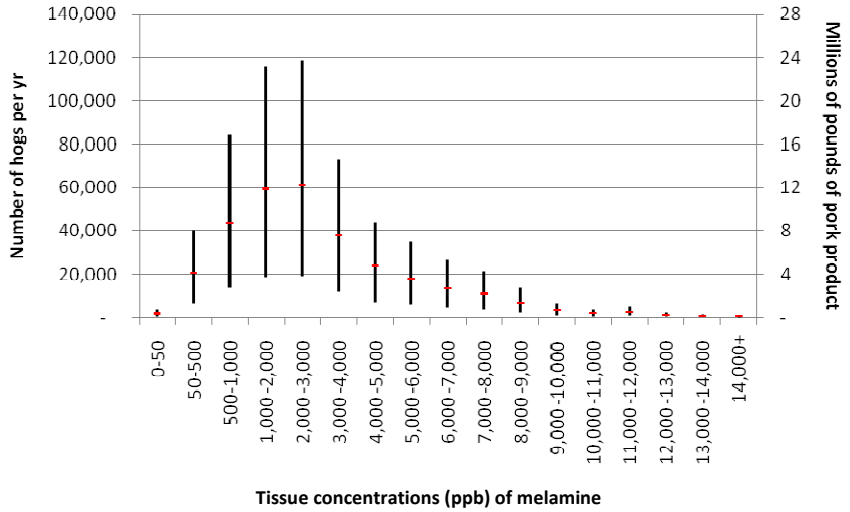

FIGURE 6-6 Distribution of number of hogs (left scale) and pounds of pork product (right scale) with various melamine tissue concentrations. A range and a median value are shown for each concentration.

TABLE 6-6 Melamine Doses to Humans Who Consume Pork with Various Melamine Tissue Concentrations[a]

Animal Tissue Concentration (ppb)	Mean Dose to Humans Who Consume Contaminated Pork (μg/kg per day)	90th Percentile Dose to Humans Who Consume Contaminated Pork (μg/kg per day)
50-500	0.10	0.50
500-1,000	0.30	1.00
1,000-2,000	0.60	2.00
2,000-3,000	1.00	3.00
3,000-4,000	1.40	4.00
4,000-5,000	1.80	5.00
5,000-6,000	2.20	6.00
6,000-7,000	2.60	7.00
7,000-8,000	3.00	8.00
8,000-9,000	3.40	9.00
9,000-10,000	3.80	10.00
10,000-11,000	4.20	11.00
11,00-12,000	4.60	12.00
12,000-13,000	5.00	13.00
13,000-14,000	5.40	14.00
14,000+	5.80	15.00

[a]Values rounded to reflect accuracy of calculations.

Personal Controllability

As in the case with infant formula, the first element necessary for personal controllability—knowledge that a potential risk exists—is lacking in the case of melamine contamination in pork products. There is no way for a consumer to detect directly that contamination exists, so individual knowledge of a potential problem is possible only after institutions detect and announce that such a problem is possible. At the time of this case study, no such determination had been made. If the possibility of contamination in pork products becomes known, consumers can eliminate the risk of melamine-contaminated food through appropriate decisions concerning shopping and food consumption. Direct consumption of pork could be avoided relatively easily. The problem is more difficult with prepared and processed foods. For some products, the pork content may be discernable only by a close examination of the food label. Nevertheless, a vigilant, knowledgeable consumer could avoid all pork. It is estimated that 80-90% of typical consumers could avoid pork products if they were aware that a contamination problem were possible.

Ability to Detect Adverse Health Effects

Detecting any problems that might occur from human consumption of pork contaminated with melamine as described in this case study would also be extremely difficult. The estimated health effects of such potential contamination are estimated to be so low as to be indistinguishable from the base-rate occurrence of kidney stones and renal disease. The lifetime rate of the occurrence of kidney stones in the United States is 5-9%, and about 2 million doctor or hospital visits per year are related to kidney stones (NIDDK 2010). Even if the effects were much higher than estimated here, detecting them and correctly attributing them to contaminated pet food used as hog feed would be very unlikely. The committee

estimated that the annual number of doctor and hospital visits for kidney stones would have to increase by 50-100% (1-2 million additional visits) before such an increase would be noticed and raise questions about what the cause might be. Given the estimated size of the exposed population (500,000, with a range of 240,000 to 1.7 million), even if all the members of the exposed population suffered from kidney stones serious enough to require medical attention, the effect is unlikely to be detected.

Ability to Mitigate Adverse Health Effects

Once a problem is detected, steps can be taken to ensure that contaminated ingredients are no longer used in animal feed. The use of pet food as a supplement is not widespread, and the producers that do use the supplement could be readily identified and notified, so mitigating risks by preventing future use of contaminated animal feed would be relatively easy. Tracking down the contaminated pork products and removing them from market shelves would be more difficult. If the contaminated hogs are mixed with unaffected hogs across a wide geographic area, massive recalls would be required for all the contaminated meat to be removed from the market. If the contaminated hogs are butchered and packaged in isolated batches, a recall could be tightly focused and very effective, but this is not likely to be the case.

In late 2008, dioxin contamination was found in pork fat in Ireland as a result of routine residue testing. In that case, the farm from which the pigs came was known, and follow-up testing quickly identified the source of the problem as contaminated feed. Within a week, all pig and cattle farms that received similar feeds had been identified. However, because they were unable to trace pork products on the market to specific farms, ultimately a full recall of all pork products was issued less than 2 weeks after contamination was found (Casey et al. 2010). The full recall minimized the chance of exposures to dioxin in pork after the problem was identified. The committee assumed that a similar scale of recall might be necessary in the United States to mitigate fully the adverse health effects of melamine in pig feed if such contamination were found.

As discussed in the case study on foods, recalls have limited effectiveness in preventing the recalled product from being consumed; in general, foods with a long shelf-life are easier to recall than those with a short shelf-life, and foods with a well-understood supply chain and those stored at centralized locations for a relatively long time are easier to recall. The distribution of pork products in the United States probably most closely resemble that of shrimp as discussed in Chapter 4. So the committee estimated that 10-50% of the potential adverse effects associated with pork already contaminated at the time that the possibility of adverse effects was recognized could be mitigated through recall.

USING THE RISK CHARACTERIZATION TO SUPPORT DECISION-MAKING

Table 6-1 summarizes the characterization of public-health consequences of potential melamine contamination of infant formula and animal feeds. It shows that, given current understanding and uncertainty, the scenario of contaminated infant formula leads to higher public-health consequences than the one of pork products contaminated from animal feed. Although the likelihood and concentration of melamine in animal feed are higher than those in infant formula, infants consume formula directly, whereas the melamine in animal feed has to be sequentially absorbed in two animal systems before posing a risk. Fewer people are likely to suffer adverse health effects from melamine contamination of animal feeds than from contaminated infant formula, but the ability to detect and mitigate such problems is lower for animal feeds than for infant formula.

The risk characterization described is an illustration of the evaluation that could be used to rank potential risks: it compares the effects of potential contamination in various products by using a consistent set of metrics. A ranking alone, however, is often not sufficient to support decisions about the next step: about what actions should be taken to reduce risks or, in this example, to understand the extent of the risks better. An extension of the analysis above could be conducted, assuming that there is an interest in

conducting laboratory tests of infant formula and animal feed to determine whether melamine contamination is present and, if so, at what levels.

For purposes of illustration, the committee assumed that resources are sufficiently limited for laboratory testing to be possible for only one of the products—either infant formula or hog feed but not both. FDA must decide which product to focus testing resources on (an example of a targeting decision). If no testing is conducted on either product, the attribute table developed above provides a summary of the expected consequences on the basis of the current estimates of the likelihood of various melamine concentrations. However, the outcomes of a decision to test either product will need to be evaluated to understand fully the implications of the alternative testing decisions.

Assume that testing will indicate either that melamine is present or that it is not present, and further assume that FDA will intervene with a program to mitigate the associated risk if melamine contamination is detected in either infant formula or hog feed. That next step is illustrated in the decision tree in Figure 6-7. The committee notes that FDA could decide to conduct a risk intervention for either or both products without testing, but for the purposes of this case study, the committee assumed that intervention would occur only after testing.

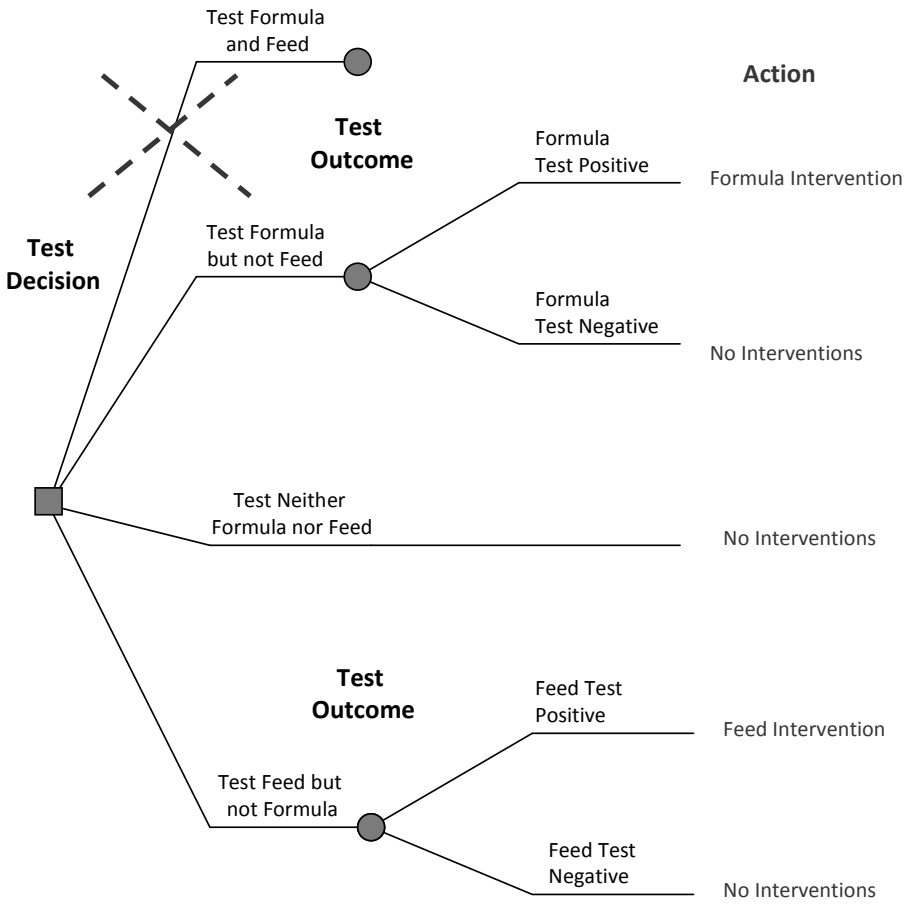

FIGURE 6-7 A decision tree for a testing-with-intervention decision involving infant formula and hog feed.

Finally, if an intervention is conducted, the net outcome depends on the effectiveness of the intervention in reducing or eliminating risks, other public-health effects caused by the intervention itself, *and* the melamine concentrations in the product not tested. Figure 6-8 depicts a single branch of the decision tree in Figure 6-7—testing infant formula only—with the outcomes delineated. Of the four outcomes, only one would require the construction of a new attribute table, "Outcome given formula intervention." In developing the attribute table, it is important to include not only the reduction in the risk of melamine-induced illnesses, but any additional health effects associated with the intervention. For example, pulling large quantities of baby formula from market shelves will greatly increase anxiety and lead to changes in how babies are fed. The public-health consequences of potential contamination of animal feed have already been evaluated in the sections described above, and the public-health consequences of having no contamination in infant formula can be assumed to be zero.

Actual testing decisions and processes clearly have additional complications. As a final step toward making this description more realistic, one can explicitly incorporate the possibility of a testing error in the decision tree. A false-positive would lead to an intervention for melamine when melamine is not present in the product; a false-negative would lead to no intervention even though melamine is present. Figure 6-9 illustrates that additional consideration for the decision option to test infant formula only.

The upper half of the tree in Figure 6-9 shows the outcomes if the test is positive: a positive test leads to an intervention. If contamination were actually present, the intervention reduces the risks from that contamination. The intervention may carry additional public-health consequences, and it is important to consider all such consequences, including potential adverse effects. For example, the health costs of a formula recall could be large. If no contamination is present, there are no "benefits" (that is, illnesses averted) to balance the increased "costs" of the intervention. The bottom half shows the outcomes if the test is negative; here, the consequences differ, potentially significantly, if the test is accurate or not. In this formulation, a type II testing error (false negative) results in a missed opportunity to mitigate a risk, including all the associated health effects.

There are clearly many decisions associated with the testing procedures that have not been discussed in this simple example. The likelihood of type I and type II errors depends on the parameters of the specific tests conducted, such as sample size and equipment used. Balancing the two types of errors (for example, test sensitivity and specificity) and costs would require an expansion of the decision tree. Ideally a fully developed decision tree with all relevant variables included could be used to complete value-of-information calculations and sensitivity analyses.

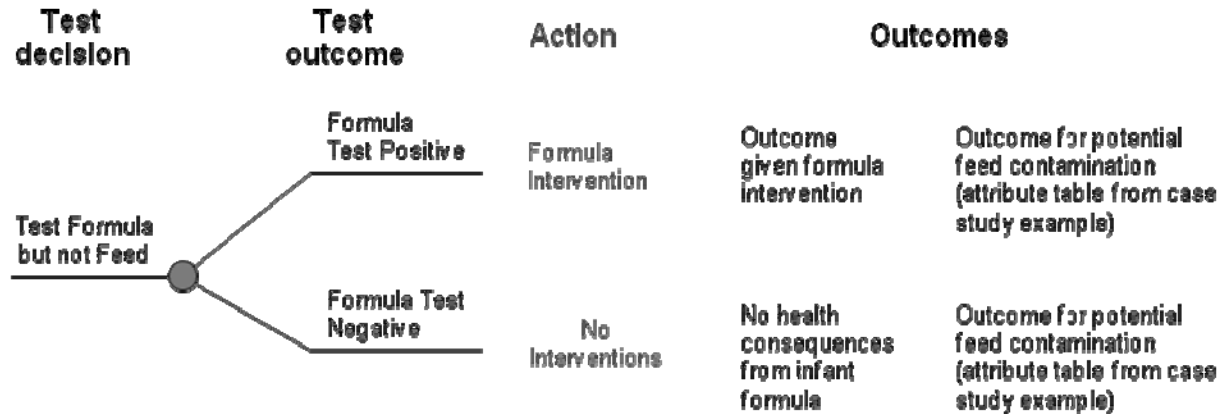

FIGURE 6-8 Potential outcomes of a decision to test infant formula but not animal feed, assuming a test without error, and interventions if the test result is positive.

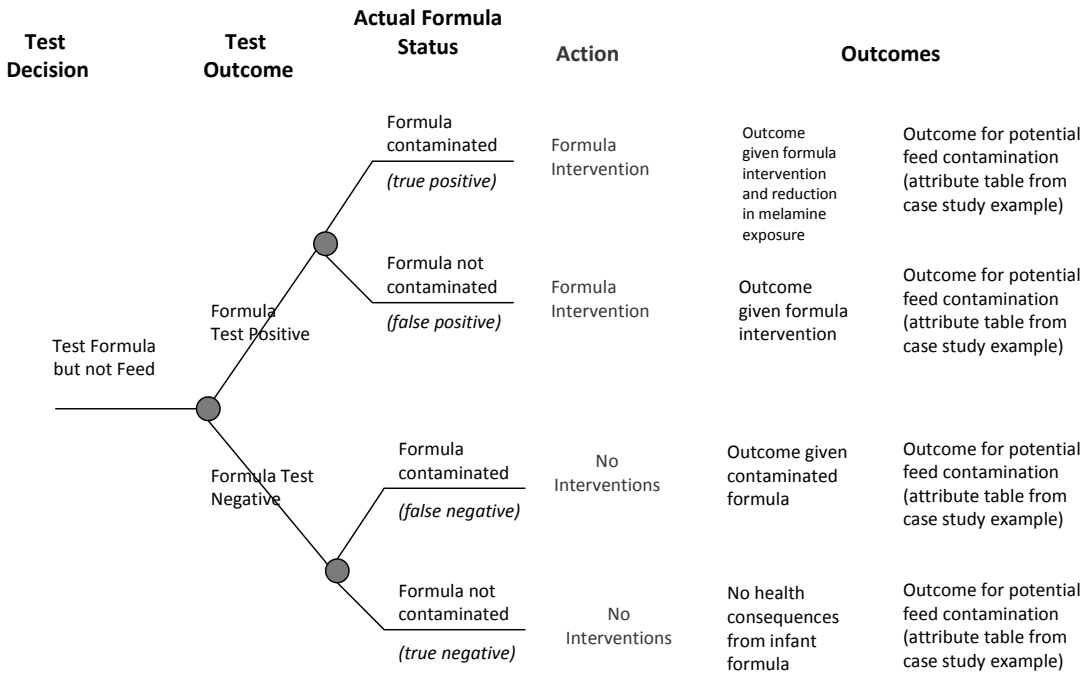

FIGURE 6-9 Tree for the decision to test formula but not feed, including uncertainty about test accuracy.

REFERENCES

AP (Associated Press). 2007. 104 Deaths Reported in Pet Food Recall. New York Times [online]. Available: http://www.nytimes.com/2007/03/28/science/28brfs-pet.html?ex=1176264000&en=8ee0fb91fd221e4b&ei=5070 [accessed March 30, 2011].

Bradley, D. 2008a. Melamine in Milk. Sciencebase.com, September 17, 2008 [online]. Available: http://www.sciencebase.com/science-blog/melamine-in-milk.html [accessed Nov. 8, 2010].

Bradley, D. 2008b. Melamine Scandal Widens. Sciencebase.com, September 22, 2008 [online]. Available: http://www.sciencebase.com/science-blog/melamine-scandal-widens.html [accessed Nov. 8, 2010].

Briefel, R.R., K. Reidy, V. Karwe, and B. Devaney. 2004. Feeding infants and toddlers study: Improvements needed in meeting infant feeding recommendations. J. Am. Diet. Assoc. 104(Suppl. 1):S31-S37.

Brown, C.A., K.S. Jeong, R.H. Poppenga, B. Puschner, D.M. Miller, A.E. Ellis, K.I. Kang, S. Sum, A.M. Cistola, and S.A. Brown. 2007. Outbreaks of renal failure associated with melamine and cyanuric acid in dogs and cats in 2004 and 2007. J. Vet. Diagn. Invest. 5(5):525-531.

Casey, D.K., J.S. Lawless, and P.G. Wall. 2010. A tale of two crises: the Belgian and Irish dioxin contamination incidents. British Food Journal. 112(10):1077-1091.

CDC (Centers for Disease Control and Prevention). 2010. Breastfeeding. Frequently Asked Questions. Centers for Disease Control and Prevention [online]. Available: http://www.cdc.gov/breastfeeding/faq/index.htm [accessed Nov.8, 2010].

CFS (Centre for Food Safety). 2008. Melamine. Frequently Asked Questions. Centre for Food Safety, Hong Kong [online]. Available: http://www.cfs.gov.hk/english/whatsnew/whatsnew_fstr/whatsnew_fstr_Test_dairy_product_FAQ.html [accessed Nov.8, 2010].

Family Education. 2010. How Much and How Often to Formula Feed. Family Education [online]. Available: http://life.familyeducation.com/formula-feeding/baby/39374.html [accessed Nov. 9, 2010].

FAO/WHO (Food and Agriculture Organization of the United Nations and World Health Organization). 2010. Proposed Draft Maximum Levels for Melamine in Food and Feed (N13-2009). CX/CF 10/4/5. Joint FAO/WHO Food Standards Program, Codex Committee on Contaminants in Foods, 4th Session, Izmir, Turkey, 26-30 April 2010 [online]. Available: http://www.cclac.org/english/committees/cf_e.php [accessed Nov. 9, 2010].

FDA (U.S. Food and Drug Administration). 2007. Interim Melamine and Analogues Safety/Risk Assessment. U.S. Food and Drug Administration [online]. Available: http://www.fda.gov/ScienceResearch/SpecialTopics/PeerReviewofScientificInformationandAssessments/ucm155012.htm [accessed Nov. 9, 2010].

FDA (U.S. Food and Drug Administration). 2008. Interim Safety and Risk Assessment of Melamine and Its Analogues in Food for Humans. U.S. Food and Drug Administration [online]. Available: http://www.fda.gov/Food/FoodSafety/FoodContaminantsAdulteration/ChemicalContaminants/Melamine/ucm164522.htm [accessed Nov. 9, 2010].

Gossner, C.M., J. Schlundt, P. Ben Embarek, S. Hird, D. Lo-Fo-Wong, J.J. Beltran, K.N. Teoh, and A. Tritscher. 2009. The melamine incident: Implications for international food and feed safety. Environ. Health Perspect. 117(2):1803-1808.

Itchmo. 2007. FDA Warns Tembec About Violations at Animal Feed Plant. Itchmo: News for Dogs and Cats [online]. Available: http://www.itchmo.com/fda-warns-tembec-about-violations-at-animal-feed-plant-3061 [accessed Nov. 9, 2010].

Lim, L.O., S.J. Scherer, K.D. Shuler, and J.P. Toth. 1990. Disposition of cyromazine in plants under environmental conditions. J. Agric. Food Chem. 38(3):860-864.

Ma, J. 2008. Adding Chemicals to Milk Common: Insiders. South China Morning Post A3, September 29, 2008.

NHF (National Hog Farmer). 2005. Hog Farm Numbers Continue to Fall. National Hog Farmer [online]. Available: http://nationalhogfarmer.com/ar/numbers-fall/ [accessed Jan. 25, 2011].

NIDDK (National Institute of Diabetes and Digestive and Kidney Diseases). 2010. Kidney and Urologic Diseases Statistics for the United States. National Kidney and Urologic Diseases Information Clearinghouse (NKUDIC), National Institute of Diabetes and Digestive and Kidney Diseases, National Institute of Health, Bethesda, MD [online]. Available: http://kidney.niddk.nih.gov/kudiseases/pubs/kustats/ [accessed Nov. 9, 2010].

Pork Checkoff. 2009. Quick Facts: The Pork Industry at a Glance. National Pork Board, Des Moines, IA [online]. Available: http://www.pork.org/MediaLibrary/FlipBooks/QuickFacts2010/document.pdf [accessed Nov. 9, 2010].

Puschner, B., R.H. Poppenga, L.J. Lowenstine, M.S. Filigenzi, and P.A. Pesavento. 2007. Assessment of melamine and cyanuric acid toxicity in cats. J. Vet. Diagn. Invest. 19(6):616-624.

Reuters. 2008. Melamine Use "Rampant" in China Feed Business. Reuters, September 25, 2008 [online]. Available: http://www.reuters.com/article/idUSPEK166968 [accessed Nov. 9, 2010].

Tembec BTLSR Inc. 2007. Re: Warning Letter CIN-DO 07-33572-18. Response letter to Carol A Heppe, Director, Cincinnati District, U.S. Food and Drug Administration, Cincinnati, OH, from Randy Fournier, President, Tembec BTLSR Inc., Toledo, OH. October 3, 2007 [online]. Available: http://www.fda.gov/ICECI/EnforcementActions/WarningLetters/2007/ucm171217.htm [accessed Nov. 9, 2010].

USDA (U.S. Department of Agriculture). 2009. Overview of the U.S. Hog Industry. National Agricultural Statistics Service, Agricultural Statistics Board, United States Department of Agriculture [online]. Available: http://usda.mannlib.cornell.edu/usda/nass/hogview//2000s/2009/hogview-10-30-2009.pdf [accessed Jan. 25. 2011].

USDA (U.S. Department of Agriculture). 2010. Quarterly Hogs and Pigs. National Agricultural Statistics Service, Agricultural Statistics Board, United States Department of Agriculture [online]. Available: http://usda.mannlib.cornell.edu/usda/nass/HogsPigs/2010s/2010/HogsPigs-12-27-2010.pdf [accessed Jan. 25, 2011].

Weise, E. and J. Schmit 2007. FDA limits Chinese food additive imports. USA Today [online]. Available: http://www.usatoday.com/money/industries/2007-04-30-chinese-imports-usat_N.htm [accessed March 30, 2011].

WHO (World Health Organization). 2008. Melamine and Cyanuric Acid: Toxicity, Preliminary Risk Assessment and Guidance on Levels in Food. World Health Organization [online]. Available: http://www.who.int/foodsafety/fs_management/Melamine.pdf [accessed Nov. 9, 2010].

WHO (World Health Organization). 2009. Toxicological and Health Aspects of Melamine and Cyanuric Acid. Report of a WHO expert meeting in collaboration with FAO supported by Health Canada Geneva: World Health Organization [online]. Available: http://www.who.int/foodsafety/publications/chem/Melamine_report09.pdf [accessed Nov. 9, 2010].

7

Conclusions and Recommendations

The committee applied its risk-characterization framework to four case studies, each of which was based on decision scenarios provided to the committee by the Food and Drug Administration (FDA). For each case study, the committee illustrated how its framework could be applied; it defined the specific decision options to be compared and developed a risk-attribute table to characterize the public-health consequences of the alternative decisions; and it provided a discussion of how the risk characterization could be used to support the specific decision options being compared. In doing so, the committee relied on members' expert judgments and data gleaned primarily from publicly available literature and databases. This chapter presents the committee's perspective on the challenges and the lessons learned from its experience in applying the framework to the case studies. General conclusions and suggestions for future directions are provided at the end of the chapter.

LESSONS LEARNED FROM THE CASE STUDIES

Value of Discussion and Multiple Points of View

The development and analysis of each case study in this report benefited greatly by involving both subject-matter expertise and decision-analysis expertise. The decision focus of the framework, specifying and comparing the outcomes of specific decisions, did not come naturally to committee members who had more detailed scientific expertise related to FDA products and product categories. They were initially inclined to look more broadly at the effects of the product being considered, and some iteration and detailed discussion were necessary to narrow the focus of discussion to the comparison of specific options. For example, in discussions about the evaluation and comparison of the risks associated with various food products (Chapter 4), the committee was initially tempted to include a discussion of the health benefits of each food product as well as the risks from each. After focusing on the decisions that the comparison might be used to support (that is, allocation of food-inspection resources), the committee determined that the focus should be on the risks and, more specifically, on the risks that could be averted by improved or more rigorous inspections. The committee notes, however, that the targeting-decision case study did not explicitly compare the health consequences of the current inspection processes with those of changes in the inspections; if that decision were to be evaluated, additional steps would be necessary.

The decision analysts on the committee were able to focus the subject-matter experts on a relatively constrained decision context, to identify the sequence of information needs, and to assist the subject-matter experts in making judgments about the array of possible effects on the basis of sparse data. The decision analysts, of course, could not provide the specialized and detailed knowledge necessary to identify and recognize the most relevant data for a specific decision context. The committee was hampered in one case study (the effects of potential melamine contamination of infant formula) by the

lack of detailed subject-matter expertise among the committee members; as a result, the committee had much less confidence in the estimates of the risk attributes of the case study than in the estimates of the other three case studies.

In all case studies, the discussions and interactions between committee members with different backgrounds and expertise were critical for the use of the risk-characterization framework. On the basis of its experience, the committee concludes that FDA will benefit from including multiple stakeholders in its decision-making process, from defining decisions to gathering information and ultimately formulating conclusions. Just as shared decision-making (Charles et al. 1997) is beneficial for medical treatment decision-making (including information-sharing and consensus-building), it will be beneficial for FDA strategic decision-making.

Defining the Decision Context

The committee found that it was critical in each case to define the decision options to be evaluated and compared clearly, so that appropriate risk information for the decision-making process could be obtained. In all cases, decision-analytic structuring was used to organize thinking about the decision context. Analytic reasoning and basic structuring tools, such as influence diagrams (see, for example, Figure 3-2), were used to identify the various factors that needed to be considered to develop estimates of the public-health consequences of the alternative decision options.

For mitigation-selection decisions, as illustrated with the vaccine-withdrawal case study in Chapter 3, defining clear and distinct decision options to be compared was straightforward. Although the example was deliberately chosen to be a simple comparison of a yes-no variety, it would have been easy to expand the set of options being considered to include more nuanced options and to make it a more complex example.

For targeting decisions, as illustrated with the evaluation of three food categories in Chapter 4 and the evaluation of melamine testing in Chapter 6, defining the decision context and the options to be compared was more complex. In fact, the food case study focused on comparing the health consequences associated with consumption of the different food categories but stopped short of evaluating different resource allocations. The evaluation and comparison presented in Chapter 4 could be used to support a risk ranking or could be used as one input into a targeting decision. For example, if FDA were deciding where to target additional food-safety inspection resources, understanding the public-health consequences as characterized in Chapter 4 would be an important input. As described in Chapter 2, for targeting decisions, the options or alternatives theoretically available to FDA are vast; virtually any amount of a resource could be allocated to the identified products or product categories and is constrained only by the total resources available. However, before substantial time and effort are invested, the many options possible need to be narrowed judiciously, and such narrowing will necessarily involve input from FDA management in addition to the technical staff.

Finally, strategic-investment decisions, as illustrated by the evaluation of enhanced surveillance of medical implants in Chapter 5, proved the most difficult to formulate and evaluate with the framework. In theory, defining the options for this case study—current surveillance compared with enhanced surveillance—was simple. In practice, however, the committee members had to speculate about the details of what an enhanced surveillance program would entail to enable them to estimate its effects. That proved to be a difficult task and one that clearly had substantial effect on the estimates derived. As described by FDA (Bertoni 2010), strategic-investment decisions are typically long-term capacity-building investments.

Characterizing the Public-Health Consequences of Each Option

For each case study, various tools were used to develop the estimates necessary to characterize the public-health consequences and populate the attribute tables shown in each case study. In some cases,

the quantities of interest could be estimated directly from available data; for others, several steps—some with considerable uncertainty—were required to generate estimates. In simple cases, exploratory descriptive statistics and bounding analysis were used. For example, estimates of the number of people exposed to the risk of foodborne illness caused by pathogens in leafy greens required an estimate of the number of people who consume leafy greens in a year, which could be based on readily available information on food consumption. In more complex cases, a series of estimates and relatively complex calculations were used to derive estimates for the attribute table. For example, estimating the number of deaths that might occur from melamine contamination of infant formula required estimates of the fraction of infant formula potentially contaminated, the concentration of melamine in that contaminated formula, the amount of contaminated formula consumed by an infant (which varied with age), the estimated dose received by infants consuming contaminated formula, and an estimate of the dose-response relationship. To develop the final estimate, the committee had to identify and structure the various factors and their relationships, estimate each of the critical factors (by using a combination of descriptive histories, bounding analyses, and judgment), and calculate the resulting numbers (in this case, using Monte Carlo simulation methods).

The risk-characterization framework provides guidance on the estimates that are necessary to compare decision options but not on how those estimates are to be developed. In working through the case studies, the committee encountered several challenges that FDA will also face in applying this framework. Some of the challenges are discussed below.

Challenges in Finding and Interpreting Data

The success of the proposed risk-characterization framework depends on the ability to populate the attribute table. Common challenges among all case studies were finding and interpreting data to support the required estimates. In the vaccine-withdrawal case study described in Chapter 3, for example, determining the excess risk of intussusception attributable to the vaccine was difficult; there were few data on the background rate of intussusceptions and little information on whether the rotavirus infection might cause intussusceptions in some cases. There was also speculation at the time that the cases of intussusceptions occurring after vaccination would have happened anyway: that is, the causal relationship between the RRV-TV vaccine and intussusception was speculative.

For the food case study described in Chapter 4, there were several data challenges. The industry segment is so large and diverse that information on volumes, producers, and distribution is not readily available. The committee chose a simple measure of the size of the exposed population (the number who consume any of the product over the course of a year) partly because more detailed data about annual consumption and consumption quantities were not readily available. An additional complication for many food categories is the lack of morbidity and mortality data. Although information exists on the estimated number of illnesses, deaths, and hospitalizations because of foodborne pathogens generally, no direct data exist on the attribution of those illnesses to specific commodities. The committee used other sources of data and made a number of assumptions to support estimates of the attribution of the illnesses to the specific food categories.

In the melamine case study described in Chapter 6, the committee was severely hampered by lack of data. The committee notes, however, that the lack of data reflects the reality of the situation being evaluated; at the time of the case study, virtually no information was available on the concentrations of melamine in various products in the United States. Regardless, the committee was able to estimate attribute values by using the available data, assumptions, and judgments and to produce a table that would have been helpful for decision-making.

The case study on strategic-investment decisions described in Chapter 5 highlighted some additional challenges. Although lack of data is clearly a problem, inaccuracies in the available data are barriers to accurate evaluations and make it difficult to identify newly emerging risks. Furthermore, when data are difficult to obtain because they are in multiple locations and in inconsistent formats, developing

the required estimates is again hampered. Thus, having data in a format that will support decision-making is clearly advantageous. For example, in the case study on strategic-investment decisions, the committee observed that a simple count of adverse event reports in the databases (MAUDE and MDR databases) does not yield a suitable estimate for determining the probability of an adverse event. There is potential for both over-reporting and under-reporting in the information contained in those databases, and there is no information on the total number of devices implanted. Furthermore, it is unclear in the reported data whether an adverse health effect suffered by a patient who has a medical device is a result of the device or is a result of some cause unrelated to the device.

Use of Expert Judgment

Expert judgment and data were inextricably intertwined in the committee's approach to each case study in this report. In some case studies, the committee did not have much direct information; in others, a large variety of data were available. In all cases, assumptions were necessary about how to interpret the data to complete the risk-attribute table. Among the case studies, evaluating the potential strategic-investment decision of enhanced postmarket surveillance of implanted medical devices proved challenging with respect to data and the need to rely more heavily on "pure" expert judgments. Those challenges arose partly from the fact that the specifics of the enhanced surveillance system had to be hypothesized, and it was not clear precisely what new information would be attained or how it might be used. As discussed in that case study, for example, it is clear that better information would reduce uncertainty in the estimated number of adverse health effects (which would reduce the range between the 5th and 95th percentiles), but it is not possible to estimate what the new range would be before collecting the information.

Some type of expert judgment is always required in evaluating and comparing the potential outcomes of different decisions. Within the risk-characterization framework, decision options are to be evaluated and compared on the basis of whatever type, quantity, and quality of data are available when the decision must be made. In some cases, detailed peer-reviewed risk analyses might be available; in other cases, one may need to rely primarily on expert judgments. That flexibility allows risk information to be considered by decision-makers for any risk-relevant decisions even if detailed quantitative risk analyses are not available. The framework provides a structured way to document the data and the associated expert judgments clearly; as the framework is used more extensively, some of the analyses and data sources used for earlier studies can be leveraged to make related studies less burdensome, although some new data and new expert judgments will probably be required.

Using the Risk Characterization to Support Decision-Making

The risk-attribute table provides a succinct comparison of the decision options that were evaluated and should be useful to decision-makers interested in understanding the key differences in the public-health consequences of those options. The comparisons alone, however, are not likely to provide all the decision-relevant information that decision-makers and policy-makers need to consider, nor are they intended to do so. The focus of the framework is to enable a comparison of the potential public-health outcomes of different decisions and to provide a common language for discussing those consequences within FDA. The committee concludes that such risk information is relevant to many FDA decisions and that clear characterization of the consequences will lead to more consistent consideration of those issues. As discussed in Chapter 2, however, the committee clearly recognizes that many other factors must be considered by FDA in its risk-management decisions.

The case studies illustrate that careful examination of the attribute table may lead to clear conclusions about the relative public-health consequences of different options, as in the mitigation-selection case study in which one option dominated the other. That will not always be the case; the

summary table may simply highlight that one option is better on some attributes but worse on other attributes than the alternative, as in the comparison of food categories. In the latter case, the FDA may ultimately want to consider more formal approaches for weighing the tradeoffs among the different risk attributes to determine which option, on balance, would be preferred in terms of public-health consequences alone, as discussed in Chapter 2.

Extending the Framework to Estimate the Value of Information

When scientists conclude that more or better information is necessary and time and resources are available to obtain that information, the risk-characterization framework can be used to highlight what type of additional information on public-health consequences would be most useful by using the decision-analytic concept of the *value of information.* As discussed in Chapter 2, new information is valuable only if it has the potential to change decisions and thus potentially improve outcomes (Clemen 1996). In a well-defined decision context, value of information can be quantified, and the cost of data collection could be compared with its value.

Two of the case studies provide some insight into the potential value of information: the case study on a strategic-investment decision and the case study on a targeting decision spanning FDA centers. The former focused specifically on evaluating and comparing the public-health consequences of two levels of information collection (the current system vs an enhanced system). The latter included discussion of an extension of the case study to one in which alternative decisions would be explicitly included in the evaluation. In the context of the decision-relevant value of an enhanced postmarket-surveillance system, many changes in decisions may result from the gathered enhanced information, including possible device recalls, revised guidelines for patient selection or patient monitoring, and different device designs. If decisions to take different actions lead to different health outcomes (or lead to other decision-relevant aspects, such as operational efficiencies, public perception and trust of FDA, or political support for FDA activities), the enhanced system will have delivered information of value. Comparing the value of that information with the costs of collecting it is outside the scope of the present committee's charge, but it could be done by FDA.

CONCLUSIONS

In developing the risk-characterization framework and conducting the case studies, the committee came to the following general conclusions:

- The committee found that framing the evaluation in a decision context was more straightforward and created an evaluation that would be more relevant for decision-making than simply conducting a risk ranking of products or product categories.
- The committee found that a risk-characterization framework could be developed with a relatively small number of attributes that are applicable within and among FDA programs. Those attributes can provide FDA with a common vocabulary for discussing risk-related decisions across centers and can be used as the basis of a consistent approach for including risk components in decision-making. There is a learning process for developing and refining the attributes, and comfort with the risk-attribute vocabulary grows over time.
- On the basis of its experience in developing the case studies, the committee found that it is possible to characterize decision options by using the risk attributes and that they could be estimated by using existing data and expert judgment. The judgments that were required were not always easy, and committee members were not always comfortable in making them, but in the end the committee concluded that the case studies would provide useful, relevant, and sufficiently accurate information to be of use to a decision-maker. The committee recognizes FDA's strong preference for "data" over "expert

judgment" for obtaining estimates or making decisions. However, it is important to recognize that when a decision must be made immediately, the committee's suggested approach can provide useful information about the public-health consequences of various options in a clear and consistent way on the basis of the best information available at the time the decision must be made.

FUTURE DIRECTIONS

As a result of its efforts to develop the risk-characterization framework and the case studies, the committee offers the following suggestions concerning implementation of the framework:

- FDA should consider using the concepts defined by the risk-characterization framework and particularly the risk attributes defined in the present report as a common language for discussing risk-related aspects of various decisions. In risk-related decisions, considering the outcomes of alternative decisions in terms of the attributes identified in the present report will begin to establish consistency in risk vocabulary throughout the agency and will build a base of understanding that will enable more detailed use of the approach for evaluating and comparing decision options in the future.
- As FDA begins to use the risk attributes and risk comparisons, such as those illustrated in the case studies for comparing decision options, it may find that some aspects of the method need to be modified. The committee believes that such modifications are entirely appropriate and that this approach should evolve to meet the agency's needs as its staff gain experience in implementation of the approach.
- In its interactions with FDA, the committee came to recognize that in many cases the agency has a substantial amount of data but that the data are not collected, organized, or accessible in a format that is useful for supporting risk-based decision-making. More focus on developing and implementing structured decision processes that are based on clearly defined risk attributes and metrics will allow the agency to improve its approaches and mechanisms for collecting information. The committee emphasizes that simply collecting more data is not necessarily the best use of resources; collecting more relevant data and organizing them so that they are useful for decision-making is the key. The committee acknowledges that new data-collection approaches and efforts will require information management and technology support.
- The committee recognizes that precise predictions of the outcomes of different decisions based on the risk attributes may be difficult to develop. Data may be lacking, and scientists may be uncomfortable in making or even unwilling to make the necessary judgments to estimate the risk attributes. However, the committee emphasizes that decisions in which risk information could be valuable are made regularly and recommends that FDA use internal or external experts who are trained in and comfortable with decision analysis, risk assessment, risk management, and specifically the assessment of uncertainties to facilitate the use of the committee's framework in its initial implementation.
- The committee recognizes that FDA will need specific expertise, training, and staffing to implement the proposed risk-characterization framework consistently. As a first step, the agency could convene a workshop to educate staff in the use of the framework and use the case studies in the committee's report as models. The agency could also provide resources to staff in various programs who have innovative ideas for implementing the framework. In addition, an intra-agency group could be formed and meet regularly to share ideas and discuss the challenges of implementing the risk-based approach.

REFERENCES

Bertoni, M.J. 2010. Opening Remarks. Presentation at the 5th Meeting on Ranking FDA Product Categories Based on Health Consequences, Phase II, February 3, 2010, Washington, DC.

Charles, C., A. Gafni, and T. Whelan. 1997. Shared decision-making in the medical encounter: What does it mean? (or it takes at least two to tango). Soc. Sci. Med. 44(5):681-692.

Clemen, R.T. 1996. Making Hard Decisions: An Introduction to Decision Analysis. Belmont, CA: Duxbury Press. 664 pp.

Appendix A

Letter Report on the Development of a Model for Ranking FDA Product Categories on the Basis of Health Risks

Committee on Ranking FDA Product Categories Based on Health Consequences

Board on Environmental Studies and Toxicology

Division on Earth and Life Studies

National Research Council

Institute of Medicine

NATIONAL RESEARCH COUNCIL *AND*
INSTITUTE OF MEDICINE
OF THE NATIONAL ACADEMIES

THE NATIONAL ACADEMIES PRESS
Washington, D.C.
www.nap.edu

THE NATIONAL ACADEMIES
Advisers to the Nation on Science, Engineering, and Medicine

National Research Council
Division on Earth and Life Studies
Board on Environmental Studies and Toxicology

500 Fifth Street, NW
Washington, DC 20001
Phone: 202 334 2347
Fax: 202 334 2752

February 17, 2009

Dr. Kathleen M. Koehler
Science Policy Analyst
Office of Science and Data Policy
U.S. Department of Health and Human Services
Hubert Humphrey Building, Room 434E
200 Independence Avenue, SW
Washington, DC 20201

Dear Dr. Koehler:

At your request, the National Academies convened the Committee on Ranking FDA Product Categories Based on Health Consequences. The committee members were selected on the basis of their expertise in food safety, health economics, medical devices, vaccine safety, pharmacoepidemiology, biostatistics, comparative risk analysis, and decision analysis.

The committee was tasked with developing and applying a conceptual model to rank product categories in FDA program areas on the basis of health risks, both positive and negative aspects (that is, the committee was to consider beneficial aspects of the product categories in the context of possible adverse health consequences). The study was divided into two phases: selection of the model (phase I) and development, refinement, and application of the model to conduct a risk ranking of FDA product categories (phase II). The committee's task is described in greater detail below. This letter report fulfills the task specified for phase I of this project.

The committee held two meetings. The first included a public session during which FDA staff and other invited experts made presentations. During that session, some indicated that a model that incorporates evaluations of interventions would be particularly valuable. The committee agrees but notes two complicating factors: evaluating baseline risks among product categories is a task of great magnitude and complexity, and it is the nature of interventions to be at the individual-product level and not the product-category or program level. Therefore, the model dictated by the committee's task is not directly applicable to "intervention" analysis and cannot be used to evaluate *strategies* to reduce risk. However, the committee acknowledges that the existence of intervention capabilities is an important measure in determining risk, and model parameters will need to capture that aspect. Given the size and complexity of the task, the committee will attempt to keep the model as simple as possible, recognizing that rough estimates of risk may be all that is possible at the product-category level.

This letter report first provides background information on comparative risk analysis. Next, it outlines the conceptual model. Considerations regarding the product categories and their attributes are provided. The report concludes with a discussion of the steps needed to refine the model and conduct a

risk-ranking exercise. (There are also several attachments: a verbatim statement of the committee's task, a committee roster and biographies, a bibliography, and acknowledgment of reviewers.) The report reflects the consensus of the committee and has been reviewed in accordance with standard National Research Council review procedures.

Sincerely,

Robert Lawrence, *Chair*
Committee on Ranking FDA Product Categories
Based on Health Consequences

LETTER REPORT ON THE DEVELOPMENT OF A MODEL FOR RANKING FDA PRODUCT CATEGORIES ON THE BASIS OF HEALTH RISKS

BACKGROUND

In 1986, U.S. Environmental Protection Agency (EPA) Administrator Lee Thomas asked 75 scientists and managers to develop a report on the "relative importance" of various environmental threats that were mainly in EPA's jurisdiction. In 1987, the group issued *Unfinished Business* (EPA 1987), which categorized environmental threats in 31 problem areas, defined largely along existing programmatic lines. The group identified and divided the risks according to four important attributes with respect to the characterization of the 31 environmental problem areas: cancer risk, noncancer risk, ecologic risk, and "welfare" effects. The resulting report provided separate assessments (ranked low, medium, or high) for each of the four attributes. A key insight of *Unfinished Business* was that EPA's resource allocations appeared to be more in line with what the public perceived as the most important risks than with the priorities identified by the agency's experts. EPA asked its Science Advisory Board (SAB) to review the report, and the SAB released a follow-up report, *Reducing Risk* (EPA SAB 1990), which endorsed the broad comparative risk analysis (CRA) approach and produced findings similar to those in *Unfinished Business*.

In addition to spawning many applications of CRA at the office, region, state, and local levels (Minard 1996; Jones 1997), the early CRA efforts led to questions about how best to facilitate comparisons and identify useful attributes for characterizing risks or risk-reduction opportunities. The EPA SAB noted that ranking risks or ranking the alternative actions that might be available for reducing risks would probably yield different rankings (EPA SAB 1990). In particular, although some risks might rank high, they might also be associated with very expensive or uncertain risk-reduction actions and therefore be unamenable to intervention according to cost-benefit criteria. In addition, if risks associated with some low-priority areas can be addressed effectively with certainty at low or no cost, their low-priority status should not prevent these "bargains" from being recognized.

Progress on CRA method development continues, although its use remains relatively limited. Finkel and Golding (1995) noted that the "comparison of risks involves values in at least five areas: defining what we mean by 'risk'; selecting the endpoints to consider; categorizing the risks for comparison; selecting a time frame for evaluating the adverse effects; and gauging the seriousness of the consequences." In February 1994, a workshop organized by Resources for the Future for the President's Office of Science and Technology Policy brought together researchers in CRA with the goal of developing a systematic process for comparing risks among different federal agencies (Davies 1996). As part of that work, researchers from Carnegie Mellon University developed a framework for ranking risks that included both quantitative and qualitative measures of relevant programmatic attributes (Fischhoff 1995; Morgan et al. 1996). They included health-impact measures (such as morbidity and mortality) and psychometric measures that research shows play an important role in the evaluation of risks (such as fairness, scientific understanding, and uncertainty). That work spawned a series of research projects and papers that refined and applied the framework (e.g., Morgan et al. 1999, 2001; Long and Fischhoff 2000; Morgan et al. 2000; DeKay et al. 2001; Florig et al. 2001; Willis et al. 2004, 2005; Fischhoff 2006; Gutiérrez et al. 2006; Bronfman et al. 2007, 2008a,b), including a discussion directly related to food safety (DeKay et al. 2005). Recently, those risk-ranking methods have been adopted by a variety of national and international entities. For example, the U.S. Army Corps of Engineers is using the methods to rank hurricane mitigation opportunities on the Louisiana Gulf Coast (USACE 2008), researchers at the University of North Carolina School of Public Health and RAND Corporation are using them to develop an environmental-health strategy and action plan for the United Arab Emirates (UNC 2008), and the British government is using them to communicate with and gather information from the general public on health-related priority-setting strategies (HM Treasury 2004, 2005a,b; OGC 2008). Regardless of the application, such projects share the goals of collecting and presenting risk information in a systematic

manner to guide and assess informed judgments. After assessment, those judgments may serve as a valuable input into a decision-making process focused on evaluating difficult policy choices.

STATEMENT OF TASK AND COMMITTEE'S APPROACH

In light of the increased use of CRA by federal agencies, the Department of Health and Human Services (DHHS) and the Food and Drug Administration (FDA) asked the National Research Council to convene an expert committee to develop and apply an evidence-based conceptual model and methods for ranking categories of products addressed by FDA programs. The conceptual model and methods were to focus on ranking product categories according to the ranges of magnitude of various potential health consequences to U.S. users of the products at the individual level and the population level, taking both adverse and beneficial effects into consideration. To accomplish its task, the committee was to include the following activities: consult with the sponsor to select FDA product categories to be ranked; consider products currently in use and near-term future products expected to come under FDA purview; review selected scientific literature bearing on adverse and beneficial health consequences; consider the scientific literature broadly to include social-science and economics literature, gray literature, and regulatory-policy literature; seek opportunities to assess health consequences in a way that allows results to be compared among broad product categories; identify information needed to address key uncertainties; assess the performance of the evidence-based model for ranking the selected product categories and identify next steps for model refinement; and where applicable and feasible, consider the potential effects on population health if risk-reduction strategies curtail the beneficial use of products.

The committee was asked first to produce a brief letter report that describes the scientific conceptual framework to be used to rank product categories (phase I) and then to perform ranking exercises by using the proposed conceptual framework (phase II). In neither phase was the committee to recommend regulatory strategies; those choices entail policy judgments that transcend scientific and technologic considerations. This letter report fulfills phase I of the project.

To accomplish its task, the committee held a public session at its first meeting, during which it heard presentations from FDA staff in the various program areas and from experts in the fields of decision analysis and CRA. The committee reviewed numerous scientific publications on CRA and literature provided by FDA. On the basis of its review and the statement of task, the committee selected a model that has the capacity to evaluate multiple product categories and compare them; to evaluate the magnitude and variation in distribution of both favorable and unfavorable effects; to improve FDA's discharge of its responsibilities as they affect public health; to evaluate new product categories, risks, benefits, and other considerations; and to include multiple non-health-related outcomes of interest, such as equity and the quality of scientific understanding. The committee recognized that the model should be able to function to the greatest extent possible with sparse information. Although the primary focus of the committee was human health, the model considers animal health and welfare to be consistent with the full scope of activities conducted by FDA.

The CRA exercise requested in the statement of task is a valuable tool in determining relative risks among product categories, but such exercises are not sufficient to guide many policy decisions unless they incorporate additional concerns. For example, the absolute risk in a category may not be a good indicator of the potential to reduce risk in that category or of the potential to reduce risk by any specific action. Measures of the potential benefits of specific actions are critical for resource-allocation decisions. Likewise, the presence of a health risk may or may not be associated with the economic costs or benefits of addressing the risk, the equity concerns (who pays and who benefits), the likelihood and timeframe of achieving the stated risk reduction, or the public's perceptions of the risk-mitigation options. To the extent that any of those concerns are or should be important in making policy decisions, the proposed CRA alone would not be sufficient for making decisions. On the basis of the guidance that the committee received from the statement of task and from clarification offered by the sponsor during its first public meeting, it concluded that discussion of the merits of other theoretical frameworks that might

be valuable in assessing risk-mitigation alternatives was outside its task. Baseline ranking of risks is a necessary but not sufficient step in the more general decision-making process.

RANKING MODEL

Conceptual Framework

On the basis of recent literature, the committee concluded that the best approach to ranking FDA product categories on the basis of health risk is to use a conceptual framework similar to that described by Florig et al. (2001). Unlike other comparative methods (such as the World Health Organization's Global Burden of Disease Study [Murray and Lopez 1996; WHO 2008], which ranks solely on the basis of utility loss from illness), the approach described by Florig and colleagues allows for disparate items, such as cosmetics and vaccines, to be ranked, as will be necessary for FDA. Furthermore, although this is not explicitly required by FDA, the selected approach is designed to accommodate qualitative and quantitative variables in the formal ranking process. That will facilitate inclusion of important variables and may greatly improve the utility of the proposed approach for FDA.

Figure 1 summarizes the two phases and the multistep process envisioned by the committee. As illustrated in Figure 1, steps A and B involve defining the FDA product categories to rank and identifying the risk attributes to describe the categories, respectively. In phase I of this project, which is summarized in this letter report, the committee has proposed a *preliminary* list of categories and attributes. The final determination of the categories and attributes will require further input from knowledgeable FDA staff, iteration, and refinement and will be completed in phase II of this project. The committee's final report will discuss the final categorization and identification of attributes and the process used to make those determinations. Step C requires describing the categories in terms of the attributes, step D involves performing the risk-ranking exercises, and step E involves summarizing and evaluating the results of the risk-ranking process. Steps C-E will be accomplished in phase II of this project, the results of which will be described in the committee's final report. Each step in the multistep process is described in greater detail in the sections that follow.

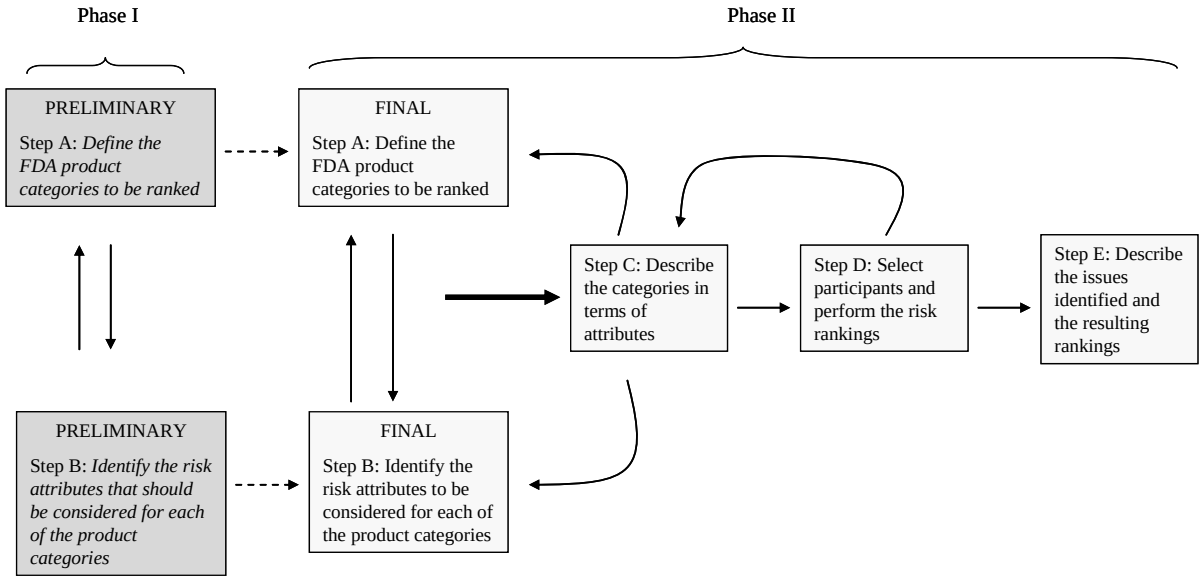

FIGURE 1 Framework of health-risk ranking model. Source: Adapted from Florig et al. 2001. Reprinted with permission; copyright 2001, *Risk Analysis.*

Uncertainties in the risk-ranking model will be captured in two ways. For some attributes, such as those measuring mortality risks, a quantitative approach can be used by providing a mean (or median) with population upper and lower bounds. Other attributes can be selected to represent uncertainties with a qualitative description. For example, an attribute that captures the quality of scientific understanding will indicate knowledge gaps, which will lead to uncertainty in the ranking exercise. The committee recognizes that successful ranking will require both iteration and further engagement with FDA.

Step A: Defining the Product Categories

Any ranking process must begin with development of a list of the risk items to rank. Many approaches exist to categorize risk, and as Morgan et al. (1996) note, "no single categorization scheme is likely to serve all … needs." Morgan et al. (2000) describe criteria for defining categories and state that categories should be "exhaustive so that no relevant risks are overlooked," "mutually exclusive so that risks are not double-counted," and "homogenous so that all risk categories can be evaluated on the same set of attributes." Furthermore, the categories should be relevant to the organizational structure, legislative mandates, and risk-management activities of the organization. Among the other criteria listed by Morgan et al. (2000) is the goal of keeping the number of categories to a number that makes the risk-ranking exercise feasible. Depending on the techniques used during the ranking, a feasible number would generally be in the range of 15-30. Given the number of products that FDA regulates, the committee recognizes that the number of categories could be expanded too much and make risk-ranking impossible. That means that the task of ranking products for FDA as a whole must of necessity focus on highly aggregated product categories.

FDA provided the committee with an initial list of product categories, which is shown in Table 1. The list includes 28 categories, and it primarily mirrors FDA's organizational structure and statutory and regulatory authorities (that is, it is broken down according to FDA's existing five product-focused centers). Although the committee noted that the categorization could alternatively focus on type or magnitude of hazard, it concluded that the final selection of the product categories for ranking will require further input from FDA staff who have specific expertise in the FDA products. Valuable input from FDA will include data on the size of each potential category (for example, with respect to numbers of regulated individual products or firms and relative market sizes in dollars). The committee expects that some of the categories will expand, others will contract, and some will be substantially revised. For example, the committee questioned the product categories suggested for medical devices. Currently, medical devices are defined according to risk and classified as class I, II, or III devices. Accordingly, this scheme may be more appropriate for categorization of medical devices. Furthermore, the committee notes that the present list includes categories at different levels of specificity, and this could obscure the value of the ranking exercise. The committee will determine whether additional categories will be needed to address near-term future products, such as nanomaterials, or whether those products can be integrated into the existing categories. Using the criteria of Morgan et al. (2000) discussed above and input from FDA, the committee will be able to determine the most appropriate product categories for evaluation in phase II of this project.

Step B: Identifying the Attributes

Ranking risks requires identifying the important attributes of the risks. Morgan et al. (1996) described criteria for selecting attributes and noted that attributes should be comprehensive, non-overlapping, stand-alone, measurable, and minimal to reduce the complexity of the risk-ranking exercise. As a preliminary scoping exercise, the committee selected five attribute groups related to exposure, severity of effect, ability to anticipate and prevent adverse events, ability to mitigate adverse events, and benefits of products or product categories. Each group contains multiple specific attributes, which are shown in Table 2. The committee emphasizes that Table 2 is only a preliminary list. Overlapping

attributes must be explicitly noted to avoid double-counting in the risk-ranking exercise, and most important, attributes must be selected that are applicable between and within the broad FDA product categories. That exercise will be challenging and will require input from FDA staff who have specific expertise in FDA product categories. The committee will use the criteria of Morgan et al. (1996) discussed above and further input from FDA to finalize the list of risk attributes in phase II of this project. At the conclusion of this process, the attributes will be clearly defined and well understood by FDA staff and the risk rankers.

The committee defined *exposure* as the condition of being subject to some effect or influence and considered the five risk attributes shown in Table 2 to be appropriate for quantifying or describing it. The *exposed population* is the percentage of the U.S. population potentially exposed. *Cumulative incidence* is the number of new cases of illness, injury, or other health-related events attributable to an exposure during a specified period in a specified population and is expressed as a rate. *Prevalence* is the number of cases of a health-related state or event that exist in a specified population at a particular time, regardless of when they began or how long they have existed, and can be expressed as a rate. *Vulnerable groups* refer to people who have increased susceptibility to adverse outcomes because of genetics, age, socioeconomic status, occupational or environmental exposure, or physiologic state; this attribute could be described in terms of the number and size of vulnerable populations. *Cluster* refers to a group of people who are at excess risk for adverse events that are related temporally, by proximity, or by source; this attribute could be described in terms of group size.

TABLE 1 Suggested Initial List of Product Categories for Ranking Provided by FDA

Program Area	Product Categories
Food, cosmetics, and dietary supplements	Food
	Produce
	Eggs and dairy
	Processed food
	Seafood
	Cosmetics
	Dietary supplements
	Food and color additives
Drugs and biologics regulated as drugs	Over-the-counter drugs
	Diagnostic prescription drugs
	Preventive prescription drugs
	Prescription drugs that are life-sustaining
	Prescription drugs for treatment for symptoms or improvement in quality of life
	Prescription drugs that are used cosmetically
Biologic products other than those regulated as drugs	Vaccines
	Blood and tissue products
	In vitro diagnostics related to donor testing
	Devices regulated as biologics
	Allergenics
	Cell and gene therapy
Veterinary products	Approved animal drugs
	Unapproved animal drugs
	Animal feeds
	Pet food
Medical devices and radiation-emitting products	Critical devices for professional use
	Noncritical devices for professional use
	Noncritical devices for lay use
	Nonmedical radiation-emitting devices
	Patient cables and lead wires

TABLE 2 Risk Attributes for Model for Ranking FDA Product Categories on the Basis of Health Risk

Attribute Groups	Risk Attributes
Exposure	Exposed population
	Cumulative incidence
	Prevalence
	Vulnerable groups
	Cluster
Severity of effect	Mortality
	Morbidity
	Vulnerable groups
	Catastrophic event
	Diffusion effects beyond intended use[a]
	Animal health
Ability to anticipate and prevent adverse events	Quality of scientific understanding
	Availability of substitutes
	History of problems and corrective actions
	Availability of quality standards, guidelines, or standard operating procedures (quality assurance and quality control, good manufacturing practices)
	Variability in product composition or performance
	Vulnerability of supply chain
Ability to mitigate adverse events	Availability of substitutes
	Availability of corrective actions
	Traceability
	Latency
	Ability to recall
	Reversibility
Product benefits	Mortality reduction
	Morbidity reduction
	Efficacy and effectiveness
	Animal welfare

[a]*Diffusion effects* is an attempt to capture effects on people who do not use the product.

Severity of the effect is described by six attributes as shown in Table 2. *Mortality* could be measured as expected number (or range of numbers) of deaths per year in the United States attributable to the product category. *Morbidity* could be quantified by one or more of the following: number of doctor visits per year, number of hospitalizations (or hospital days) per year, number of lost work days per year, total cost of treatment, and number of chronic cases per year. Metrics need to account for acute and chronic illnesses. The committee discussed using quality-adjusted life years (QALYs) to capture the differences but concluded that the method might be difficult to apply to something as heterogeneous or complex as the product categories given the expected paucity of data. However, if data are available, QALYs should be included in the risk-ranking model. *Vulnerable groups* could be represented by the percentage of all deaths that occur in vulnerable groups or the percentage of deaths in a vulnerable group that are attributable to the product category. *Catastrophic event* refers to a low-probability event with the potential for a severe outcome and could be characterized by the number of deaths in a worst-case scenario. The attribute *diffusion effect* is an attempt to capture effects on people who do not use the product (for example, transgenerational effects resulting from product use); this attribute could be characterized by a minor-moderate-major significance descriptor. *Animal health* could be measured as the number of animals that die from use of animal products; food animals and companion animals could be evaluated separately.

The committee considered six specific attributes to characterize the ability to anticipate and prevent adverse events. *Quality of scientific understanding* is related to product knowledge, that is,

background information available on a product, the number of scientific studies, and the quality of the studies. The *availability of substitutes* attribute is related to how critical a product is. The attribute *history of problems and corrective actions* reflects the availability of information on such events as adverse reactions and manufacturing defects and whether actions have been needed to address problems. The attribute *availability of quality standards, guidelines, or standard operating procedures* attempts to capture the idea that manufacturing products in accordance with standards, guidelines, or standard operating procedures improves FDA's ability to control product quality and minimize adverse events (for example, the availability of practices to produce uniform, high-quality products and the limits of the practices). *Variability in product composition or performance* refers to inherent differences in products that result even when good manufacturing practices are adhered to or that result from intended or unintended deviations from good manufacturing practices, including inadvertent contamination, equipment malfunction, or deliberate adulteration. *Variability in performance* is also related to person-person differences in response to a product, such as the side effects experienced after taking a drug; such variability may be related to individual characteristics or to how the product is used. Finally, *vulnerability of the supply chain* refers to the potential disruption in manufacture or distribution of a product if there are any problems with obtaining inputs to any step of the process. All the risk attributes in this group could be defined by a binary yes-no descriptor or a high-moderate-low descriptor.

Six attributes are used to describe the ability to mitigate adverse events. *Availability of substitutes* and *corrective actions* would provide methods to mitigate an adverse event by replacing a product with an equivalent or similar product, by modifying manufacturing or distribution processes, or by communicating with consumers about potentially harmful products. *Traceability* is the ability to identify the sources of all product components, affected products, or components in a supply chain, or people potentially exposed to products. *Latency* is defined as the interval between the first exposure and the observation of an adverse event; long latency would make it difficult to anticipate and prevent adverse events. *Ability to recall* is the ability to remove from the supply chain a product that is identified as carrying unacceptable risks. *Reversibility* is the amelioration of adverse events. Each of the attributes could be defined with a yes-no descriptor or a high-moderate-low descriptor.

The final attribute group attempts to incorporate the beneficial aspects of products to the general and target populations. Some benefits are linked to protecting the population from adverse consequences; some are linked to diagnosing, treating, or preventing disease; and others are linked to promoting health, such as those related to nutrition. The benefits may be best quantified by expected reduction in mortality (that is, the number of lives saved per year in the United States from intended use of the product category) and by expected reduction in morbidity (for example, the number of disabilities or of hospitalizations avoided per year in the United States). The *efficacy and effectiveness* attribute captures how well the current program is working (for example, Is it cost-effective?). *Animal welfare* tries to capture the idea that food production animals are important food sources and contribute substantial benefits to the U.S. economy and that companion animals enhance quality of life and may provide some health benefits. Products that enhance the quality of life of the animals would provide a benefit to society. A high-moderate-low descriptor could be used to characterize that attribute. The committee emphasizes that capturing beneficial aspects of product categories will be a challenging task. Some benefits are obvious and often discussed, and others are less well known, especially if products have long-term consequences. Many benefits will be difficult to quantify or rank, such as those related generally to improving quality of life or social values by providing people with a variety of choices or with novel products.

Because it is highly desirable to have an attribute table that allows easy side-by-side comparisons, there should be fewer than 20, and ideally fewer than 15, attributes. Winnowing the attribute list down will require merging some of the attributes shown in Table 2. For example, the "Ability to anticipate and prevent adverse events" set, which currently has six specific attributes, will probably contain only one or two in the final version. If those six attributes were all described by using binary variables, 64 combinations would be possible. Given that only 15-30 product categories will be ranked, that number of attributes overspecifies the ranking task. Winnowing of the attribute list will occur naturally as overlapping attributes are highlighted and preferred attributes are selected. In some cases, entire

attributes groups could be eliminated (for example, is the "exposure" group necessary if the "severity of effect" group contains mortality and morbidity rates and counts?).

The committee notes that for each product category the risk rankers will receive a summary "pamphlet" that will include the attribute table (which will display all the attribute values in a consistent format compatible with easy side-by-side comparisons) and supporting information that helps to put the attribute values in context. Because the attributes are selected to be universal and applicable between product categories, there are relevant details that cannot be captured in a simple listing of quantified attributes. For example, the amplifying text could include descriptions of the vulnerable groups relevant for the product category, the extent to which life-cycle calculations were completed or limited, or animal impacts that were not explicitly shown in the attribute table. The supplemental information will be concise and organized in a standard manner so that comparisons between product categories are possible.

Step C: Describing the Product Categories by Using the Attributes

In this initial phase of its work, the committee has not reached formal recommendations for the follow-on steps in the framework, but it recognizes the difficulties ahead, particularly in completing step C, in which the values of each attribute (step B) must be described for each product category (step A). In some ranking exercises, the items for ranking are unique and narrowly defined. In this case, however, the broad product categories, such as vaccines, contain many individual products, each with their own attribute values. Thus, aggregation will present important challenges with respect to attribute characterization. For some attributes, such as those measuring mortality risks, a summation across all items in the product category may capture the relevant values, with upper and lower bounds on the quantified values providing an indication of the uncertainties in the calculations. However, because of the heterogeneity of some of the product categories, difficulty will arise in assigning unique values for many attributes. It may be desirable to split large product categories to improve homogeneity of ranking (for example, vaccines might be divided into killed vaccines and live vaccines). For some attributes, aggregation may have the effect of making them nonvariant among the set of items being ranked. That problem could be mitigated by focusing the attributes in question on the mechanisms that lead to the greatest concern for morbidity and mortality. Thus defined, the attribute would distinguish product categories with well-established cause-effect linkages from categories that are less well understood.

The necessary refinement of the attribute definitions will not be possible until the committee can re-engage the necessary FDA personnel through follow-on workshops in phase II of this project. Discussions could lead to a regrouping of the product categories and attribute description. The process by which the first three steps are completed will generate valuable information. In fact, the committee emphasizes that the process that it undertakes with FDA is likely to be a more important tool and provide more insights than the list of ranked product categories alone. Creating robust attributes that are systematically and accurately measured across the wide array of FDA regulatory responsibilities will focus agency thinking and help to communicate the considerations underlying agency decision-making to a broad range of stakeholders.

Steps D and E: Conducting the Ranking and Analyzing the Results

Once the attributes are fully defined for each product category, the ranking will begin, probably with at least two approaches: a holistic ranking based on the rankers' overall preferences and a ranking based on application of a formal mutiattribute model. The committee anticipates selecting the ranking procedures after further discussion with FDA to learn how it plans to use the rankings.

In a holistic ranking, the preferences for, importance of, and trade-offs between the attributes are not made explicit by the rankers. The rankers carefully review the summary material and, using their own judgment, rank the product categories. Guidance as to how to structure the ranking procedure can make the task more manageable. For example, guidance provided could be to complete a rough sorting into

three preliminary categories—high, moderate, and low risk—before completing the series of pairwise comparisons. In contrast, the formal multiattribute model approach requires the rankers to consider each attribute and their range of values and to state explicitly the relative trade-offs between attributes through the elicitation of attribute weights. A model is built for each ranker by using those weights, and a ranking of the product categories is calculated. Sensitivity analyses are performed on the ranks to determine the influence of the assessed weights and the value functions.

Although the two basic approaches to complete the rankings differ substantially, there are advantages in using both approaches in a single exercise (Palmgren et al. 2000). With a holistic approach, overlapping attributes can be considered (that is, the combination of attribute values may provide insights into relevant details); however, with a formal multiattribute model, double counting is not allowed, and attribute weight assessment must reflect this requirement. The results of using one approach can improve the understanding of the other, and ultimately a revised ranking informed by the results of both is likely to be preferred by the rankers. Analyses conducted after the ranking can determine the relative importance of the two approaches in generating a ranker's final ordering. Developing a satisfactory multiattribute model for the ranking task has many potential benefits, including the ability to add new or revised product categories and to see quickly how they fit among the categories already ranked. Once again, understanding how FDA will use the ranking results will determine the preferred approaches. The rankings could be used internally by senior administrators as input into their own strategic planning or as a way to capture public perceptions and communicate policy to the general public. Both applications offer merit but would certainly dictate different approaches to documentation and communication of the process. Regardless of the specific uses, determining the level of agreement among the rankers could provide valuable insights that are not now available to FDA.

NEXT STEPS

The committee has proposed a conceptual model for ranking FDA product categories and has suggested preliminary categories and risk attributes. The committee emphasizes that participation of FDA staff in each program area is essential for the development of a successful and useful model. Development of the model will require a series of three workshops. The first workshop will involve discussion of the model's product categories and attributes that will lead to refinement, revision, and adoption of both for further use in steps C-E. This iterative cycling between product categories and attribute definitions will ensure that the follow-on steps can be completed. In addition, the engagement of FDA personnel will serve as an educational opportunity so that the model, its application, and its limitations can be understood and appreciated. The second workshop will involve populating the model matrix with data to allow the ranking exercises to be performed and will require much effort on the part of FDA personnel who will be responsible for gathering data and providing input values. Populating the model matrix will almost certainly require the determination of values that have not been estimated previously and rely on the use of subjective expert judgment. As indicated above, the difficulty in providing values will probably vary considerably among product categories. The third workshop will involve conduct of the actual ranking exercises. The committee emphasizes that the development of the model is an iterative process as reflected in Figure 1. Findings from any of the three workshops may necessitate adjustments and refinements of earlier steps in the model. The committee's final report will summarize and evaluate the outcome of the workshops and provide recommendations for using the risk-ranking model as an input in a decision-making process.

The committee notes that the risks and benefits vary substantially between and within product categories, and that will pose a challenge in developing and implementing the ranking process. However, that challenge makes the ranking exercise and resulting ranking valuable for FDA because unaided comparisons also face the challenge of comparing apparently incomparable product categories but without the common metrics that the committee will be recommending. Therefore, the ranking exercise is a logical first step for FDA. The committee provides an outside perspective on the challenge of

comparing risk among disparate categories, and its recommendations, with input from FDA, will help to identify a framework for making advances in FDA management processes and decisions.

Attachments:
A Statement of Task
B Committee Membership
C Bibliography
D Acknowledgment of Reviewers

Attachment A

STATEMENT OF TASK

An expert committee will develop and apply an evidence-based conceptual model and methods to rank product categories within the broad types of products addressed by programs of the U.S. Food and Drug Administration (FDA). The conceptual model and methods will focus on ranking product categories according to the potential ranges of magnitude of various health consequences to U.S. users of the products at individual and population levels, taking both adverse and beneficial effects into consideration. The committee will begin by selecting, in consultation with DHHS and FDA, categories of products within FDA mandates for human and veterinary drugs, biologics, medical devices, foods, cosmetics, and products that emit radiation. The committee will then review selected scientific literature bearing on adverse and beneficial health consequences related to these product categories. It will develop a scientific conceptual framework for potential use in guiding product category rankings based on expert judgments and related analysis of the types and potential ranges of magnitude of health consequences to U.S. users of the products (phase I). Using this framework, the committee will perform ranking exercises through expert elicitation and analysis or other appropriate methods (phase II).

In carrying out its task, the committee will include the following activities:

- In selecting product categories for ranking, consider products currently in use and near-term future products expected to come under FDA purview.

- Seek opportunities to assess health consequences in a way that allows results to be compared across broad product categories.

- Where data or assessment methods are deficient for evaluating a product category, identify information needs for addressing key uncertainties and present evaluations.

- Assess the performance of the evidence-based model for ranking the selected product categories and identify next steps for further refinement of the model.

- In assessing health consequences, consider both the risks and the beneficial aspects of product use, and where applicable and feasible, consider the potential impact on population health if beneficial product use is curtailed through risk reduction strategies.

- In reviewing selected scientific literature, the committee shall consider the scientific literature broadly, to include, as appropriate, social science and economic literature, grey literature, and regulatory policy literature.

The committee will not recommend regulatory strategies, because those choices will entail policy judgments that transcend scientific and technologic considerations.

Seven months after initiation of the study, the committee will prepare a brief letter report describing the conceptual model and methods it will use to rank product categories in its final report.

Attachment B

COMMITTEE MEMBERSHIP

ROBERT LAWRENCE, *Chair*, Johns Hopkins University, Baltimore, MD
JAMES ANDERSON, Case Western Reserve University, Cleveland, OH
FRANCISCO DIEZ-GONZALEZ, University of Minnesota, St. Paul
KATHRYN EDWARDS, Vanderbilt University, School of Medicine, Nashville, TN
SUSAN ELLENBERG, University of Pennsylvania, Philadelphia
PAUL FISCHBECK, Carnegie Mellon University, Pittsburgh, PA
HELEN JENSEN, Iowa State University, Ames
ROBIN KELLER, University of California, Irvine
DAVID MELTZER, University of Chicago, Chicago, IL
SANFORD MILLER, University of Maryland, College Park
RICHARD PLATT, Harvard Medical School, Boston, MA
KIMBERLY THOMPSON, Harvard School of Public Health, Boston, MA

STAFF

ELLEN MANTUS, Project Director
DAVID A. BUTLER, Senior Program Officer
NORMAN GROSSBLATT, Senior Editor
HEIDI MURRAY-SMITH, Research Associate
PANOLA GOLSON, Senior Program Assistant

BIOGRAPHIES

Robert S. Lawrence (IOM), *Chair*, is the Center for a Livable Future (CLF) professor and director of the CLF in the Department of Environmental Health Sciences, professor of health policy and international health at the Johns Hopkins Bloomberg School of Public Health and professor of medicine at the Johns Hopkins University School of Medicine. His expertise and research interests include community and social medicine, human rights, health promotion and disease prevention, evidence-based decision rules for prevention policy, and food security. Dr. Lawrence is a master of the American College of Physicians and a fellow of the American College of Preventive Medicine. He is a member of the Institute of Medicine and has served on numerous National Academies committees, most recently the Committee on Adolescent Health Care Services and Models of Care for Treatment, Prevention, and Health Development and the Committee to Evaluate Measures of Health Benefits for Environmental, Health, and Safety Regulation. Dr. Lawrence received his MD from Harvard Medical School and trained in internal medicine at the Massachusetts General Hospital.

James M. Anderson (IOM) is professor of pathology, macromolecular science, and biomedical engineering at Case Western Reserve University. His research interests range from his activity as a pathologist in clinical implant retrieval and evaluation to fundamental mechanistic studies focused on tissue, cell, and blood interactions with biomaterials. Dr. Anderson is the recipient of the Elsevier Biomaterials Gold Medal for the most significant contributions to biomaterials science from 1980 to 2005 and the Society of Investigative Pathology Chugai Mentoring Award. He has been involved in the International Standards Organization Task Force to Develop Standards for Medical Device Safety for the last 18 years. He is editor-in-chief of the *Journal of Biomedical Materials Research*. Dr. Anderson is a

member of the Institute of Medicine and has served as a member of its Committee on Postmarket Surveillance of Pediatric Medical Devices and Committee on Capturing the Full Power of Biomaterials for Military Medical Needs. He received his MD from the Case School of Medicine and his PhD in chemistry from Oregon State University.

Francisco Diez-Gonzalez is an associate professor in the Department of Food Science and Nutrition at the University of Minnesota. His research expertise is in food-safety microbiology, foodborne pathogens, safety of fresh fruits and vegetables, preharvest control of pathogenic *E. coli*, bioterrorism agents, and safety of organic food. Dr. Diez-Gonzalez teaches courses in food safety and food microbiology. He has served on the University of Minnesota Institutional Biosafety Committee, and he has advised both undergraduate and graduate students. He is also the recipient of the New Career Excellence Award for the College of Human Ecology at the University of Minnesota. He is member of the Editorial Board of the *Journal of Food Protection* and the *Journal of Food Analytical Methods*. Dr. Diez-Gonzalez received his PhD in food science from Cornell University.

Kathryn M. Edwards (IOM) is Sarah H. Sell Chair in Pediatrics and the director of the Vanderbilt Vaccine Research Program at Vanderbilt University Medical Center. Her research focuses on the evaluation of vaccines for the prevention of infectious diseases in adults and children. She is a fellow of the Infectious Diseases Society of America and of the American Academy of Pediatrics. Dr. Edwards has served as a member of the Advisory Committee on Immunization Practices of the Centers for Disease Control and Prevention and the Vaccines and Related Biological Products Advisory Committee of the Food and Drug Administration. She has also served as a member of the National Academies Committee to Assess the Safety and Efficacy of the Anthrax Vaccine. Dr. Edwards received her MD from the University of Iowa College of Medicine.

Susan S. Ellenberg is professor of biostatistics and associate dean for clinical research at the University of Pennsylvania School of Medicine. Her research focuses on the design and analysis of clinical trials and the assessment of medical-product safety. Dr. Ellenberg is associate editor of *Clinical Trials* and of the *Journal of the National Cancer Institute*. She is a fellow of the American Statistical Association, the Society for Clinical Trials, and the American Association for the Advancement of Science. She has served as a member of the National Academies Planning Committee for the IOM Drug Safety Report: Resource Implications, Committee on the Assessment of the U.S. Drug Safety System, and Committee on Applied and Theoretical Statistics. Dr. Ellenberg received her PhD in mathematical statistics from George Washington University.

Paul S. Fischbeck is professor of social and decision sciences, professor of engineering and public policy, and director of the Center for the Study and Improvement of Regulation at Carnegie Mellon University. His research focuses on the quantification and communication of uncertainty, including theoretical improvements in decision analysis and numerous applied real-world problems. Dr. Fischbeck has written extensively on various applications of decision and risk-analysis methods and has won several awards from the Institute for Operations Research and the Management Sciences. He is a member of the National Research Council Marine Board and has served on several committees, including the Committee on Marine Salvage Response Capability: A Workshop and the Committee on Risk Assessment and Management of Marine Systems. Dr. Fischbeck received a PhD in industrial engineering and engineering management from Stanford University.

Helen H. Jensen is a professor of economics and head of the Food and Nutrition Policy Division of the Center for Agricultural and Rural Development at Iowa State University. Her research fields are food and nutrition policy, analysis of food-consumption behavior, economics of food safety, and health risk assessment. Dr. Jensen is on the Board of Directors of the American Agricultural Economics Association and of the Council on Food, Agricultural and Resource Economics and has recently been on the Editorial

Boards of *Food Economics*, *Agricultural Economics*, and *Agribusiness: An International Journal*. She has served on U.S. Department of Agriculture expert review panels, including the Panel on Measuring Food Security in the United States and the Panel on the Health Eating Index. She has served on several National Academies committees and is currently involved with the Committee on Nutrition Standards for National School Lunch and Breakfast Programs and the Committee on Economic Development and Current Status of the Sheep Industry in the United States. Dr. Jensen received her PhD in agricultural economics from the University of Wisconsin-Madison.

L. Robin Keller is a professor of operations and decision technologies at the University of California, Irvine. Her research is in decision analysis, risk analysis, creative problem-structuring, and behavioral decision theory. She is the editor-in-chief of *Decision Analysis*. Dr. Keller has served as program director for the Decision, Risk, and Management Science Program of the U.S. National Science Foundation, and she has conducted studies funded by the U.S. Environmental Protection Agency and the Department of Energy. She has served as a member of the National Research Council Committee to Assess the Distribution and Administration of Potassium Iodide in the Event of a Nuclear Incident, and she is currently a member of the U.S. National Committee for the International Institute for Applied Systems Analysis. Dr. Keller received her PhD from the University of California, Los Angeles.

David O. Meltzer is an associate professor in the Department of Medicine, chief of the Section of Hospital Medicine, and an associate faculty member of the Harris School and the Department of Economics at the University of Chicago. He is also director of the Center for Health and the Social Sciences and co-director of the Program on Outcomes Research Training. Dr. Meltzer's research explores problems in health economics and public policy, with a focus on theoretical foundations of medical cost-effectiveness analysis and the effects of managed care and medical specialization on the cost and quality of care. He is the recipient of numerous awards, including the National Institutes of Health Medical Scientist Training Program Fellowship, the National Science Foundation Graduate Fellowship in Economics, and the Lee Lusted Prize of the Society for Medical Decision Making, of which he is the immediate past president. Dr. Meltzer has served on several National Academies committees, most recently the Committee on the Assessment of the U.S. Drug Safety System and the Committee on Establishing a National Cord Blood Stem Cell Bank Program. He received his MD and his PhD in economics from the University of Chicago.

Sanford A. Miller is a senior fellow at the Joint Institute for Food Safety and Applied Nutrition at the University of Maryland. He was named professor and dean emeritus of the Graduate School of Biomedical Sciences at the University of Texas Health Science Center in December 2000 after serving as dean from 1987 to 2000. He is a former director of the Center for Food Safety and Applied Nutrition in the Food and Drug Administration. Dr. Miller has served on many national and international government and professional-society advisory committees, including the National Advisory Environmental Health Sciences Council of the National Institutes of Health and the Joint World Health Organization-United Nations Food and Agricultural Organization Expert Advisory Panel on Food Safety. He is a member of the National Academies Food and Nutrition Board and the Committee on Use of Dietary Supplements by Military Personnel. Dr. Miller received his PhD in physiology and biochemistry from Rutgers, The State University of New Jersey, New Brunswick.

Richard Platt is professor and chair of the Department of Ambulatory Care and Prevention and a professor of medicine at Harvard Medical School. His research focuses on the safety and effectiveness of marketed drugs and vaccines and on infectious diseases in the community and hospital settings. Dr. Platt is a former chair of the Food and Drug Administration Drug Safety and Risk Management Advisory Committee. He is a member of the Advisory Panel for Research of the Association of American Medical Colleges and has chaired the Executive Committee of the HMO Research Network, the Epidemiology and Disease Control Study Section of the National Institutes of Health, and the Steering Committee of the

Centers for Disease Control and Prevention Office of Health Care Partnerships. He has served on several National Academies committees and is a member of the Roundtable on Evidence-Based Medicine. Dr. Platt received his MD from Harvard Medical School.

Kimberly M. Thompson is associate professor of risk analysis and decision science at the Harvard School of Public Health. Her research is related to developing and applying quantitative methods for risk assessment and risk management and the public-policy implications of including uncertainty and variability in risk characterization. She has served on several National Academies committees, including the Committee for the Study of a Motor Vehicle Rollover Rating System and the Subcommittee to Update the 1999 Arsenic Report. She is a member of the Board on Environmental Studies and Toxicology. Dr. Thompson received her ScD in environmental health from the Harvard School of Public Health.

Attachment C

REFERENCES

Bronfman, N.C., L.A. Cifuentes, M.L. DeKay, and H.H. Willis. 2007. Accounting for variation in the explanatory power of the psychometric paradigm: The effects of aggregation and focus. J. Risk Res. 10(4):527-554.

Bronfman, N.C., L.A. Cifuentes, and V.V. Gutierrez. 2008a. Participant-focused analysis: Explanatory power of the classic psychometric paradigm in risk perception. J. Risk Res. 11(6):735-753.

Bronfman, N.C., E.L. Vazquez, V.V. Gutierrez, and LA. Cifuentes. 2008b. Trust, acceptance and knowledge of technological and environmental hazards in Chile. J. Risk Res. 11(6):755-773.

Davies, J.C. 1996. Comparing Environmental Risks: Tools for Setting Government Priorities. Washington, DC: Resources for the Future.

DeKay, M.L., H.K. Florig, P.S. Fischbeck, M.G. Morgan, K.M. Morgan, B. Fischhoff, and K.E. Jenni. 2001. The use of public risk ranking in regulatory development. Pp. 208-230 in Improving Regulation: Cases in Environment, Health, and Safety, P.S. Fischbeck, and R.S. Farrow, eds. Washington, DC: Resources for the Future.

DeKay, M.L., P.S. Fischbeck, H.K. Florig, M.G. Morgan, K.M. Morgan, B. Fischhoff, and K.E. Jenni. 2005. Judgment-based risk ranking for food safety. Pp. 198-226 in Toward Safer Food: Perspectives on Risk and Priority Setting, S.A. Hoffmann, and M.R. Taylor, eds. Washington, DC: Resources for the Future.

EPA (U.S. Environmental Protection Agency). 1987. Unfinished Business: A Comparative Assessment of Environmental Problems. EPA/230/2-87/025. Office of Policy, Planning, and Evaluation, U.S. Environmental Protection Agency, Washington, DC.

EPA SAB (U.S. Environmental Protection Agency Science Advisory Board). 1990. Reducing Risk: Setting Priorities and Strategies for Environmental Protection. SAB-EC-90-021. U.S. Environmental Protection Agency Science Advisory Board, Washington, DC [online]. Available: http://yosemite.epa.gov/sab/sabproduct.nsf/28704D9C420FCBC1852573360053C692/$File/REDUCING+RISK++++++++++EC-90-021_90021_5-11-1995_204.pdf [accessed Dec. 2, 2008].

Finkel, A.M., and D. Golding. 1995. Worst Things First? The Debate over Risk-Based National Environmental Priorities. Washington, DC: Resources for the Future.

Fischhoff, B. 1995. Ranking risks. Risk Health Saf. Environ. 6:189-200.

Fischhoff, B. 2006. Cognitive processes in stated preference methods. Pp. 937-968 in Handbook of Environmental Economics, Vol. 2. Valuing Environmental Changes, K.G. Mäler, and J.R. Vincent, eds. Amsterdam: Elsevier.

Florig, H.K., M.G. Morgan, K.M. Morgan, K.E. Jenni, B. Fischhoff, P.S. Fischbeck, and M.L. DeKay. 2001. A deliberative method for ranking risks (I): Overview and test-bed development. Risk Anal. 21(5): 913-921.

Gutiérrez, V.V., L.A. Cifuentes, and N.C. Bronfman. 2006. The influence of information delivery on risk ranking by lay people. J. Risk Res. 9(6):641-655.

HM Treasury. 2004. The Orange Book: Management of Risk-Principles and Concepts. London, UK: Her Majesty's Stationery Office. October 2004 [online]. Available: http://www.who.int/management/general/risk/managementofrisk.pdf [accessed Dec. 2, 2008].

HM Treasury. 2005a. Managing Risks to the Public: Appraisal Guidance. London, UK: Her Majesty's Stationery Office. June 2005 [online]. Available: http://www.hm-treasury.gov.uk/d/Managing_risks_to_the_public.pdf [accessed Dec. 2, 2008].

HM Treasury. 2005b. Principles of Managing Risks to the Public. HM Treasury Cabinet Office, London, UK [online]. Available: http://www.hm-treasury.gov.uk/d/risk_principles_180903.pdf [accessed Dec. 2, 2008].

Jones, K. 1997. A Retrospective on Ten Years of Comparative Risk. Prepared for American Industrial Health Council, Washington, DC, by Green Mountain Institute for Environmental Democracy, Montpelier,

VT. January 24, 1997 [online]. Available: http://heartland.temp.siteexecutive.com/pdf/23157i.pdf [accessed Dec. 2, 2008].

Long, J., and B. Fischhoff. 2000. Setting risk priorities: A formal model. Risk Anal. 20(3):339-352.

Minard, R.A. 1996. Comparative risk assessment and the States: History, politics and results. Pp. 23-62 in Comparing Environmental Risks: Tools for Setting Government Priorities, J.C. Davies, ed. Washington, DC: Resources for the Future.

Morgan, K.M., M.L. DeKay, and P.S. Fischbeck. 1999. A multi-attribute approach to risk prioritization. Risk Policy Report 6(6):38-40.

Morgan, M.G., H.K. Florig, M.L. DeKay, and P. Fischbeck. 2000. Categorizing risks for risk ranking. Risk Anal. 20(1):49-58.

Morgan, K.M., M.L. DeKay, P.S. Fischbeck, M.G. Morgan, B. Fischhoff, and H.K. Florig. 2001. A deliberative method for ranking risks (II): Evaluation of validity and agreement among managers. Risk Anal. 21(5):923-937.

Morgan, M.G., B. Fischhoff, L. Lave, and P. Fischbeck. 1996. A proposal for ranking risks within federal agencies. Pp. 111-148 in Comparing Environmental Risks: Tools for Setting Government Priorities, J.C. Davies, ed. Washington, DC: Resources for the Future.

Morgan, M.G., H.K. Florig, M.L. DeKay, and P. Fischbeck. 2000. Categorizing risks for risk ranking. Risk Anal. 20(1):49-58.

Murray, C.J.L., and A.D. Lopez, eds. 1996. The Global Burden of Disease: A Comprehensive Assessment of Mortality and Disability from Diseases, Injuries and Risk Factors in 1990 and Projected to 2020. Cambridge, MA: Harvard University Press.

OGC (Office of Government Commerce). 2008. Management of Risk (M_o_R). Office of Government Commerce, HM Treasury, London, UK [online]. Available: http://www.ogc.gov.uk/guidance_management_of_risk_4441.asp [accessed Dec. 2, 2008].

Palmgren, C.R., M.L. DeKay, P.S. Fischbeck, B. Fischhoff, and M.G. Morgan. 2000. Evaluating a Risk-Ranking Methodology. Society of Risk Analysis Annual Meeting Applications of Risk Analysis in Industry and Government, December 3-6, 2000, Washington, DC.

UNC. 2008. UNC School of Public Health to help UAE assess environmental health risks. UNC Gillings School of Global Public Health. School of Public Health News: June 9, 2008 [online]. Available: http://www.sph.unc.edu/school_of_public_health_news/unc_school_of_public_health_to_help_uae_assess_environmental_health_risks_7546_1957.html [accessed Dec. 2, 2008].

USACE (US Army Corps of Engineers). 2008. Risk-Informed Decision Framework Appendix, Draft. Louisiana Coastal Protection and Restoration Technical Report. U.S. Army Corps of Engineers, New Orleans District, Mississippi Valley Division. February 2008 [online]. Available: http://lacpr.usace.army.mil/\Report\Draft Appendices\Risk Informed Decision Framework Appendix.pdf [accessed Dec. 2, 2008].

WHO (World Health Organization). 2008. The Global Burden of Disease: 2004 Update. Geneva: World Health Organization [online]. Available: http://www.who.int/healthinfo/global_burden_disease/GBD_report_2004update_full.pdf [accessed Jan. 9, 2009].

Willis, H.H., M.L. DeKay, M.G. Morgan, H.K. Florig, and P.S. Fischbeck. 2004. Ecological risk ranking: Development and evaluation of a method for improving public participation in environmental decision making. Risk Anal. 24(2):363-378.

Willis, H.H., M.L. DeKay, B. Fischhoff, and M.G. Morgan. 2005. Aggregate, disaggregate, and hybrid analyses of ecological risk perceptions. Risk Anal. 25(2):405-428.

Attachment D

ACKNOWLEDGMENT OF REVIEWERS

This report has been reviewed in draft form by persons chosen for their diverse perspectives and technical expertise in accordance with procedures approved by the Report Review Committee of the National Research Council. The purposes of the independent review are to provide candid and critical comments that will assist the institution in making its published report as sound as possible and to ensure that the report meets institutional standards of objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process. We thank the following for their review of this report: Ann Bostrom, University of Washington; Paul Citron, Medtronic, Inc.; Alan M. Garber, Stanford University; Sandra A. Hoffman, Resources for the Future; Robert E. Johnston, University of North Carolina, Chapel Hill; Ralph L. Keeney, Duke University; Harley W. Moon, Iowa State University; Joseph V. Rodricks, ENVIRON; and Brian L. Strom, University of Pennsylvania School of Medicine.

Although the reviewers listed above have provided many constructive comments and suggestions, they were not asked to endorse the conclusions or recommendations, nor did they see the final draft of the report before its release. The review of the report was overseen by the review coordinator, Lauren Zeise, California Environmental Protection Agency, and the review monitor, John C. Bailar, III, University of Chicago. Appointed by the National Research Council, they were responsible for making certain that an independent examination of the report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of the report rests entirely with the committee and the institution.

Appendix B

Statement of Task for the Committee on Ranking FDA Product Categories Based on Health Consequences, Phase II

This statement of task pertains to the second phase of a two-phase study. In this phase, an expert committee will refine and test an evidence-based conceptual model to evaluate products and/or product categories addressed by programs of the U.S. Food and Drug Administration (FDA), excluding the tobacco program. The conceptual model and methods will have the capability of evaluating products and/or product categories according to the potential ranges of magnitude of various health consequences to U.S. users of the products at individual and population levels, taking both adverse and beneficial effects into consideration. The committee will consult with FDA to finalize products and/or product categories that are within FDA mandates for human and veterinary drugs, biologics, medical devices, foods, cosmetics, and products that emit radiation. The committee will continue to review selected scientific literature and consult with FDA to determine adverse and beneficial health consequences related to the selected products and/or product categories. It will refine the scientific conceptual model (including specification of attributes) from Phase I and will illustrate potential applications of the model in scenarios developed jointly with the FDA.

In carrying out its task, the committee will include the following activities:

- In selecting products and/or product categories for evaluation, consider products currently in use and near-term future products expected to come under FDA purview (excluding tobacco and tobacco-related products).
- Seek opportunities to assess health consequences in a way that allows results to be compared within and across broad program areas.
- Where data or assessment methods are deficient for evaluating a product and/or product category, identify information needs for addressing key uncertainties and present evaluations.
- Assess the performance of the evidence-based model for evaluating the selected product categories and identify next steps for further refinement of the model.
- In assessing health consequences, consider both the risks and the beneficial aspects of product use, and where applicable and feasible, consider the potential impact on population health if beneficial product use is curtailed through risk reduction strategies.
- In reviewing selected scientific literature, the committee shall consider the scientific literature broadly, to include, as appropriate, social science and economic literature, grey literature, and regulatory policy literature.

The committee will not recommend specific FDA activities or regulatory strategies, because those choices will entail policy judgments that transcend scientific and technologic considerations.

The committee has prepared a brief letter report that describes the conceptual model and methods.

Appendix C

Scenarios Provided by the U.S. Food and Drug Administration to the Committee on Ranking FDA Product Categories Based on Health Consequences, Phase II

In Phase I of this study, the NAS study Committee completed initial development of a conceptual model and method that has the capability of evaluating FDA-regulated products and/or product categories (excluding tobacco) according to the potential ranges of magnitude of various health consequences to U.S. users of the products at individual and population levels, taking both adverse and beneficial effects into consideration. The purpose of the study is to develop models and methods to inform the public health dimensions of an array of FDA risk-based decisions. (This project does not address regulatory strategies, because those choices will also entail legal and policy judgments that transcend scientific and technologic considerations.)

In Phase II, the NAS Committee will, among other things, use concrete examples of risk-based decisions faced by FDA to refine the model. The following scenarios describe different types of decisions that FDA faces every day. Today, FDA uses best available data and information on public health risks and benefits to inform the public health component of these decisions. The agency is asking the NAS to suggest approaches to assessing health consequences in a way that allows public health impacts to be compared within and across broad program areas, and to identify information needs for addressing key uncertainties.

1) VACCINES RISK MANAGEMENT – Many vaccines provide critical public health benefits, so that significant disruptions or interruptions in supply may have broad consequences for public health. Suppliers, supplies, and production capacity of some important vaccines are limited. Whenever concerns arise about manufacturing processes, decisions for action must weigh the relative potential consequences for safety versus the potential consequences of decreased availability. Loss of public confidence in these products and any resulting reduction in inoculation rates represents a potentially serious public health hazard. What measures of public health impacts should be taken into account in making decisions about safety interventions and/or risk management strategies, and setting priorities for action? What information is needed to support such decisions?

2) DRUG MANUFACTURING SURVEILLANCE (1) – FDA must decide how to apportion resources between surveillance of manufacturers of normally safe, over-the-counter children's cough medicine and manufacturers of a narrow therapeutic range, sterile-injectable product. The former has a nonzero risk of economically-motivated adulteration with a potentially lethal sweetener. How can we take

all relevant public health impacts into account in making such decisions about surveillance priorities and risk management strategies in this scenario? What data would inform such decisions?

3) DRUG MANUFACTURING SURVEILLANCE (2) – In allocating inspection resources, FDA must choose between engaging in a domestic drug surveillance inspection designated as high-risk by the Center versus a foreign drug inspection. What measures of public health risks and benefits should be taken into account in making decisions about drug inspection priorities, and what analyses could support such decisions? What data would inform such decisions?

4) UNPLANNED EMERGENCY ACTIVITY VERSUS HIGH-PRIORITY PLANNED INSPECTIONS – In order to accomplish two Class 1 (significant health hazard) recalls on firms that required full follow-up, FDA had to pull resources from surveillance of high-risk firms, thus jeopardizing accomplishing the targets set by the Center for surveillance of high-risk firms. Similarly, in the midst of a risk-based assignment for a Center, FDA requested four recalls; we had to marshal our workforce to address this massive undertaking. How should we analyze these situations to better understand the full array of public health impacts of these decisions? What data is needed for such analyses?

5) SETTING PRIORITIES FOR LIMITED INVESTIGATIONAL RESOURCES – An animal producer with multiple drug residue violations appears to have another violation before FDA has finished investigating the previous ones. This occurs in part because other higher priority food borne illness investigations or performance goal work take precedence in tapping the same human resources. What risks and benefits should be considered in analyzing the comparative public health impacts of follow-up violations versus other work, and how might they be compared? What data would inform such analyses?

6) COMPARATIVE RISKS/BENEFITS OF FEED SAFETY ACTITIVIES – FDA has committed significant funds for cooperative agreements with the States for work on BSE issues. A number of people have argued that there are more "risky" hazards than BSE. Now that these cooperative agreements are expiring, we have the opportunity to re-allocate these animal feed safety funds to have the biggest public health benefit. We have a number of feed safety programs: feed contaminants (e.g., heavy metals, pesticides, mycotoxins, dioxins/pcbs, pathogens); medicated feed facilities (licensed and unlicensed), drug residues in food animal tissues, and BSE. What measures of public health risks and benefits should we take into account, and how can they be compared? Are there public health measures that factor in feasibility (not just "science-based" factors), so we can get the biggest public health "bang for our buck?"

7) PRIORITY INSPECTION VERSUS CONSUMER COMPLAINT FOLLOW-UP – PDUFA (pre-approval human drug), ADUFA (pre-approval animal drug) and BIMO (clinical trial inspections related to pre-approval) assignments and foreign drug and device manufacturing inspections are designated as priority work by the Centers. However, each FDA field office must also prioritize locally whether an investigator will conduct one of these inspections or follow-up on a consumer complaint. If the complaint involves an illness, what measures of public health impact would inform the decision about whether the consumer complaint should take priority? What data would inform the analysis?

8) CRISIS MANAGEMENT VERSUS CONSUMER COMPLAINT FOLLOW-UP – During the initial melamine contamination investigation, FDA labs were flooded with food products to screen for melamine and melamine by-products. During that time, several consumer complaint samples were received that were linked to human illnesses. Those complaint samples required use of the same laboratory equipment needed for the melamine analysis. How can we use measures of potential public health impact to inform the decision about prioritizing the use of lab resources? What data is needed for such assessments?

9) CLINICAL TRIAL INSPECTION – FDA must apportion clinical trial inspection resources between trials of two drugs that, by preliminary information, are likely to be equivalently efficacious and safe products. Trial participants for the first product are collected in several large clinical trial research centers. The second product relies on clinical trial information from ~200 separate sites. There are insufficient resources to cover more than 10 sites. How can FDA incorporate public health impacts into its decisions about trial inspection priorities and risk management strategies in this scenario? What data would be needed?

10) BLOOD PRODUCTS RISK MANAGEMENT – There are high public expectations for the safety of most blood products, so that strong efforts to sustain the highest possible degrees of safety are the rule. However, because many blood products are life-saving, there is always a need to consider safety versus availability locally and nationally. In addition, there are few suppliers for some products and, the possibility of local/national shortages due to infectious disease outbreaks, manufacturing problems, or distribution problems. Selected types of blood products can present high risks but can also be life-saving/sustaining for specific groups of patients. FDA may also need to consider high profile, potential risks. As a consequence, FDA staff frequently perform ongoing risk assessments to inform actions and help to define data collection needs, safety interventions and/or risk management strategies. What measures of public health risks and benefits should be taken into account in making decisions about safety interventions and/or risk management strategies, and setting priorities for action? What data would inform such decisions?

11) DRUG RECALL – FDA must decide whether to recall a high-demand topical product that has an identified low risk of causing infection only in immunocompromised patients. We could provide consumer alerts with no recall, recall only new shipments, or also recall old product from store shelves if the product is very near expiration. How can FDA take into account the public health impacts of these options, in identifying risk management strategies in this potential recall scenario? What data would inform such decisions?

12) DOMESTIC FOOD INSPECTIONS – FDA must decide how to allocate its domestic field inspection resources between the various categories of foods, for example, between seafood and produce or between low acid canned foods and dairy products. A field work plan is developed that allocates these resources and an attempt is made to link the priorities to risk. Not all firms can be inspected each year, so priority has to be given to some firms at the expense of others and yet even the lowest risk firms should be inspected at some frequency. What are the elements of public health risk and benefit that should be taken into account when allocating these resources? How should these elements be weighted?

13) IMPORT FOOD SAMPLE COLLECTION AND ANALYSIS – FDA must decide how to allocate its import field sample collection and analysis resources between the various categories and countries of origin of foods, for example between spinach from Mexico and canned tuna from Thailand or between spice from India and soft cheese from France. In addition, consideration must be given to the hazard for which the product should be analyzed, for example, E. coli., histamine, Salmonella, and L. monocytogenes. Priorities are set in electronic screening criteria and also refined by the import inspector in the district into which the product is imported. Only a small percentage of products can be collected for analysis, so priority has to be given to some products, shippers, and countries of origin instead of others. What are the elements of public health risk and benefit that should be taken into account when making these decisions? How should these elements be weighted?

14) MEDICAL DEVICE POST-MARKET SURVEILLANCE – FDA reviews medical device adverse event reports to understand post-market device performance and to identify problems. Manufacturers' understanding of and compliance with adverse event reporting is uneven. What data could be used to determine the public health risks of inaccurate reporting and identify high risk areas? What

information could FDA consider in deciding how to focus educational resources on improving reporting for specific devices or specific manufacturers?

15) MEDICAL DEVICE RISK WEB PAGES – FDA has a number of web pages to communicate device safety information to the public. There are communications about emerging high risk issues, chronic risk areas and general safety information. What risk factors should be taken into account in setting priorities for which web pages should be updated or created first, to address the most pressing public health issues?

16) MEDICAL DEVICE CONSENSUS STANDARD PRIORITIES – FDA collaborates with standards setting organizations (IEEE, ISO, AAMI, etc) on standards that include specific acceptance criteria for relevant performance characteristics of medical devices; such standards may also include test methods appropriate for evaluating a medical device. These standards influence medical device design and testing decisions. The number of such organizations and meetings is beyond what FDA can reasonably cover. (For example, in 2008, FDA representatives attended more than 60 meetings in the US and abroad in 16 different medical specialty areas; even so, not all medical specialty areas were covered.). How can we consider medical device risk when deciding which standards to collaborate in? What factors are most likely to affect public health?

Appendix D

Biographic Information on the Committee on Ranking FDA Product Categories Based on Health Consequences, Phase II

Robert S. Lawrence (IOM), Chair, is the Center for a Livable Future (CLF) professor and director of the CLF in the Department of Environmental Health Sciences, professor of health policy and international health in the Johns Hopkins Bloomberg School of Public Health, and professor of medicine in the Johns Hopkins University School of Medicine. His expertise and research interests include community and social medicine, human rights, health promotion and disease prevention, evidence-based decision rules for prevention policy, and food security. Dr. Lawrence is a master of the American College of Physicians and a fellow of the American College of Preventive Medicine. He is a member of the Institute of Medicine and has served on numerous National Academies committees, most recently the Committee on Adolescent Health Care Services and Models of Care for Treatment, Prevention, and Health Development and the Committee to Evaluate Measures of Health Benefits for Environmental, Health, and Safety Regulation. Dr. Lawrence received his MD from Harvard Medical School and trained in internal medicine at the Massachusetts General Hospital.

Jeffery B. Bender is an associate professor of veterinary public health in the College of Veterinary Medicine and has an adjunct appoint in the School of Public Health of the University of Minnesota. He serves as the director of the Center for Animal Health and Food Safety. His main research interests are zoonoses and emerging diseases, food safety, and antimicrobial resistance. Dr. Bender teaches courses on veterinary public health, diseases common to humans and animals, factors in the emergence of zoonotic diseases, methods for epidemiologic investigations, and infectious disease surveillance. He has published numerous scientific articles and two book chapters, and he is a diplomate of the American College of Veterinary Preventive Medicine. Dr. Bender received his DVM from the University of Minnesota.

Francisco Diez-Gonzalez is an associate professor in the Department of Food Science and Nutrition of the University of Minnesota. His research expertise is in food-safety microbiology, foodborne pathogens, safety of fresh fruits and vegetables, preharvest control of pathogenic *E. coli*, bioterrorism agents, and safety of organic food. Dr. Diez-Gonzalez teaches courses in food safety and food microbiology. He has served on the University of Minnesota Institutional Biosafety Committee, and he has advised undergraduate and graduate students. He is the author of more than 50 peer-reviewed publications and 10 book chapters. He is a member of the Editorial Boards of the *Journal of Food Protection* and the *Journal of Food Analytical Methods*. Dr. Diez-Gonzalez received his PhD in food science from Cornell University.

Kathryn M. Edwards (IOM) is Sarah H. Sell Chair in Pediatrics and the director of the Vanderbilt Vaccine Research Program of Vanderbilt University Medical Center. Her research focuses on the evaluation of vaccines for the prevention of infectious diseases in adults and children. She is a fellow of the Infectious Diseases Society of America and the American Academy of Pediatrics. Dr. Edwards has served as a member of the Advisory Committee on Immunization Practices of the Centers for Disease Control and Prevention and the Vaccines and Related Biological Products Advisory Committee of the Food and Drug Administration. She is a member of the Institute of Medicine, a member of the National Academies Committee on Special Immunizations Program for Laboratory Personnel Engaged in Research on Countermeasures for Selected Agents, and a past member of the National Academies Committee to Assess the Safety and Efficacy of the Anthrax Vaccine. Dr. Edwards received her MD from the University of Iowa College of Medicine.

Susan S. Ellenberg is professor of biostatistics and associate dean for clinical research in the University of Pennsylvania School of Medicine. Her research focuses on the design and analysis of clinical trials and the assessment of medical-product safety. Dr. Ellenberg is associate editor of *Clinical Trials* and of the *Journal of the National Cancer Institute*. She is a fellow of the American Statistical Association, the Society for Clinical Trials, and the American Association for the Advancement of Science. She has served as a member of the National Academies Planning Committee for the IOM Drug Safety Report: Resource Implications, Committee on the Assessment of the U.S. Drug Safety System, and Committee on Applied and Theoretical Statistics. Dr. Ellenberg received her PhD in mathematical statistics from George Washington University.

Paul S. Fischbeck is professor of social and decision sciences, professor of engineering and public policy, and director of the Center for the Study and Improvement of Regulation in Carnegie Mellon University. His research focuses on the quantification and communication of uncertainty, including theoretical improvements in decision analysis and numerous applied real-world problems. Dr. Fischbeck has written extensively on various applications of decision and risk-analysis methods and has won several awards from the Institute for Operations Research and the Management Sciences. He is a member of the National Research Council Marine Board and has served on several committees, including the Committee on Marine Salvage Response Capability: A Workshop and the Committee on Risk Assessment and Management of Marine Systems. Dr. Fischbeck received his PhD in industrial engineering and engineering management from Stanford University.

Karen E. Jenni is president of Insight Decisions LLC, focusing on the application of decision analysis methods to energy and environmental policy issues. Much of that work emphasizes the assessment, quantification, and modeling of a variety of risks and leads to recommendations about productive areas for applied research and effective risk-management strategies. Recent public-sector projects include participation in a probabilistic volcanic-hazard analysis for the Yucca Mountain region and a multidisciplinary study of a selenium mobilization from large-scale ground disturbances, transport through environmental media, and potential effects on biota. She served as a consultant to the National Academies committee that produced *Prospective Evaluation of Applied Energy Research and Development at DOE (Phase Two)*. Dr. Jenni earned her PhD in engineering and public policy from Carnegie Mellon University.

Helen H. Jensen is a professor of economics and head of the Food and Nutrition Policy Division of the Center for Agricultural and Rural Development of Iowa State University. Her research fields are food and nutrition policy, analysis of food-consumption behavior, economics of food safety and hazard control. She has been a member of the Board of Directors of the Agricultural and Applied Economics Association and the American Council on Consumer Interests and has served on the editorial boards of a number of professional journals. Dr. Jensen serves or has served on several National Academies committees, including the Institute of Medicine Committee to Review Child and Adult Care Food Program and Meal

Requirements and the National Research Council Committee on Assessing the Nation's Framework for Addressing Animal Diseases, the Committee on Biological Threats to Agricultural Plants and Animals, the Board on Agriculture Panel on Animal Health and Veterinary Medicine, and the Committee on National Statistics Panel to Review USDA's Measurement of Food Insecurity and Hunger. She is a member of the World Health Organization Initiative to Estimate the Global Burden of Foodborne Diseases, Foodborne Disease Burden Epidemiology Reference Group. Dr. Jensen received her PhD in agricultural economics from the University of Wisconsin-Madison.

L. Robin Keller is a professor of operations and decision technologies at the University of California, Irvine. Her research is in decision analysis, risk analysis, creative problem-structuring, and behavioral decision theory. She is the editor-in-chief of *Decision Analysis*. Dr. Keller has served as program director for the Decision, Risk, and Management Science Program of the U.S. National Science Foundation, and she has conducted studies with the U.S. Environmental Protection Agency and the Department of Energy. She has served as a member of the National Research Council Committee to Assess the Distribution and Administration of Potassium Iodide in the Event of a Nuclear Incident, and she is a member of the U.S. National Committee for the International Institute for Applied Systems Analysis. Dr. Keller received her PhD from the University of California, Los Angeles.

James D. Mckean is an extension veterinarian and professor in the Department of Veterinary Diagnostic and Production Animal Medicine and associate director of the Iowa Pork Industry Center at Iowa State University. His interests include the assessment and prevention of chemical and drug residues in feed and food animals, quality assurance, food safety, food law, pseudorabies control, and animal welfare and well-being. At Iowa State University's Food Safety Consortium, his research involves sulfamethazine depletion in market-weight swine, medical and feed management for sulfamethazine, and comparing sulfamethazine depletion with sulfamethazine activity from previous experiments. Dr. McKean has previously served on several national committees on government policy development, including the Swine Futures Team, the Task Force on the Future of FSIS Veterinarians, and, as chair, the American Association of Swine Veterinarians Pork Safety Committee. He has also served as a member of the National Research Council Committee on Implications of Dioxin in the Food Supply. Dr. McKean earned his JD from Drake University and his DVM from the University of Illinois.

David O. Meltzer is an associate professor in the Department of Medicine, chief of the Section of Hospital Medicine, and an associate faculty member of the Harris School and the Department of Economics of the University of Chicago. He is also director of the Center for Health and the Social Sciences and codirector of the Program on Outcomes Research Training. Dr. Meltzer's research explores problems in health economics and public policy with a focus on theoretical foundations of medical cost-effectiveness analysis and the effects of managed care and medical specialization on the cost and quality of care. He is the recipient of numerous awards, including the National Institutes of Health Medical Scientist Training Program Fellowship, the National Science Foundation Graduate Fellowship in Economics, and the Lee Lusted Prize of the Society for Medical Decision Making, of which he is the immediate past president. Dr. Meltzer has served on several National Academies committees, most recently the Committee on the Assessment of the U.S. Drug Safety System and the Committee on Establishing a National Cord Blood Stem Cell Bank Program. He received his MD and his PhD in economics from the University of Chicago.

Sanford A. Miller is a senior fellow of the Joint Institute for Food Safety and Applied Nutrition of the University of Maryland. He was named professor and dean emeritus of the Graduate School of Biomedical Sciences of the University of Texas Health Science Center in December 2000 after serving as dean from 1987 to 2000. He is a former director of the Center for Food Safety and Applied Nutrition of the Food and Drug Administration. Dr. Miller has served on many national and international government and professional-society advisory committees, including the National Advisory Environmental Health

Sciences Council of the National Institutes of Health and the Joint World Health Organization-United Nations Food and Agriculture Organization Expert Advisory Panel on Food Safety. He is a member of the National Academies Food and Nutrition Board and the Committee on Use of Dietary Supplements by Military Personnel. Dr. Miller received his PhD in physiology and biochemistry from Rutgers, the State University of New Jersey, New Brunswick.

Richard Platt is professor and chair of the Department of Population Medicine at the Harvard Pilgrim Health Care Institute and Harvard Medical School. His research focuses on the safety and effectiveness of marketed drugs and vaccines and on infectious diseases in the community and hospital settings. Dr. Platt is a former chair of the Food and Drug Administration Drug Safety and Risk Management Advisory Committee. He is a member of the Advisory Panel for Research of the Association of American Medical Colleges and has chaired the Executive Committee of the HMO Research Network, the Epidemiology and Disease Control Study Section of the National Institutes of Health, and the Steering Committee of the Centers for Disease Control and Prevention Office of Health Care Partnerships. He has served on several National Academies committees and is a member of the Roundtable on Evidence-Based Medicine. Dr. Platt received his MD from Harvard Medical School.

John T. Watson is a professor of bioengineering at the University of California, San Diego. Formerly, he was director of clinical and molecular medicine at the National Heart, Lung, and Blood Institute of the National Institutes of Health (NIH) and contributed a total of 27 years of service to NIH. Dr. Watson is a mechanical engineer and physiologist who is interested in finding ways to reduce the time that it takes medical technology to move from a concept into the clinic. The subjects of his research include heart failure, medical-implant design and science, and biomaterials. Dr. Watson is a founding fellow and former president of the American Institute for Medical and Biological Engineering and has received numerous honors, including membership on the Greatest Engineering Achievements of the 20th Century selection committee and invitations to be a member of the Japanese Kyoto Prize nominating committee and the National Academy of Engineering Draper Prize selection committee. He is a member of the National Academy of Engineering and has participated in several National Research Council activities, including being a speaker at a Workshop on Innovation and Invention in Medical Devices and a member of the Committee to Develop a Research Agenda for Test Methods and Models to Simulate Accelerated Aging of Infrastructure Materials. Dr. Watson earned his PhD from the University of Texas Southwestern Medical Center at Dallas.

Appendix E

Factors Hypothesized as Important in Understanding Risk

The following is a list of some characteristics of hazards that have been hypothesized as important in understanding risk. They were compiled from the following sources: Lowrance (1976); Rowe (1977); Vlek and Stallen (1980); Hohenemser et al. (1983); Litai et al. (1983); Slovic et al. (1984), Jenni (1997).

Ability to contain adverse effects
Ambiguity about probability of harm
Ambiguity about severity of effects
Availability of options or alternatives
Avoidability
Blame assignable
Caused by humans
Common vs dreaded hazards
Confidence in decision-making strategies
Confidence in experts or regulators
Consequences foreseeable
Continuous vs occasional exposure
Controllability (institutional)
Controllability (personal)
Delay or timing of effects
Distribution of effects (general population vs sensitive groups)
Ease of change or correction
Ease of escape from harm
Familiarity vs Newness
Frequency of accidents
Importance of intended benefits (for example, necessity vs luxury)
Knowledge about risks and benefits
Natural vs manmade causes
Nonhuman impacts
Number of people affected
Occupational vs nonoccupational exposure
Personal awareness of risk or danger
Personal experience and knowledge
Personal influence or responsibility
Possibility of error

Preventable
Probability or frequency of adverse events or effects
Recurrence
Reversibility of effects
Severity of effect (for example, major vs minor, large vs small, fatal vs survivable, painful vs painless)
Size of the population at risk
Spatial distribution of risks
Speed with which adverse events occur
Transgenerational effects
Voluntariness

REFERENCES

Hohenemser, C., R.W. Kates, and P. Slovic. 1983. The nature of technological hazard. Science 220 (4495):378-384.

Jenni, K.E. 1997. Attributes for Risk Evaluation. Ph.D. Dissertation, Carnegie Mellon University, Pittsburgh, PA.

Litai, D., D. Lanning, and N. Rasmussen. 1983. The public perception of risk. Pp. 213-224 in The Analysis of Actual Versus Perceived Risks , V. Covello, W. Flamm, J. Rodericks and R. Tardiff, eds. New York: Plenum Press.

Lowrance, W.W. 1976. Of Acceptable Risk: Science and the Determination of Safety. Los Altos, CA: W. Kaufmann.

Rowe, W.D. 1977. An Anatomy of Risk. New York: Wiley.

Slovic, P., B. Fischhoff, and S. Lichtenstein. 1984. Behavioral decision theory perspectives on risk and safety. Acta Psychologica 56:183-203.

Vlek, C. and P.J. Stallen. 1980. Rational and personal aspects of risk. Acta Psychologica 45:273-300.