Lessons for Research Synthesis from Developing a Global Expert Elicitation

Sandra Hoffmann (USDA, Economic Research Service)

Working Paper prepared for:

Methods for Research Synthesis: A Cross-Disciplinary Workshop

Harvard Center for Risk Analysis

October 3, 2013

*Corresponding author: shoffmann@ers.usda.gov

Disclaimer: The findings and conclusions of this paper are those of the author and do not imply endorsement by any component of Harvard University or other sponsors of this workshop. The views expressed herein also do not necessarily reflect the views of the Economic Research Service or the U.S. Department of Agriculture. Comments should be directed to the author.

Acknowledgements: Tine Häld, Arie Havelaar, and other members of the WHO Foodborne Disease Burden Epidemiology Reference Group (FERG) Source Attribution Task Force, Willy Aspinall, and Roger Cooke contributed to this paper.
Lessons for Research Synthesis from Developing a Global Expert Elicitation

Sandra Hoffmann*

USDA, Economic Research Service

August 2013

In 2006, the WHO convened an effort to assess the global burden of foodborne disease. Foodborne disease is one of the last major classes of disease to be included in the WHO’s global burden of disease initiative. One of the primary reasons for this delay has been the difficulty of attributing the burden of these diseases to routes of exposure. The WHO committee organizing this research determined that it needed to rely on multiple research synthesis methods to accomplish this task. This paper focuses on the specific challenges of using expert elicitation to develop globally comparable risk attribution estimates. The paper begins by laying out data and research synthesis principles adopted in the GBD Study Protocol. It then briefly describes the specific data challenges faced in the Global Burden of Foodborne Disease Study (GBFD) and the ways research synthesis methods were used in the study as a whole. It then turns to the use of expert elicitation in the GBFD study; first looking at the criteria used to decide when to use expert elicitation rather than other research synthesis methods and then turning to the decisions that were made in developing a global expert elicitation. From this experience it draws insights into the four focus questions for this workshop:

1) What criteria should be used to evaluate the applicability of different research synthesis methods to particular types of problems and data?
2) What particular characteristics of the problem and data make the research synthesis method(s) you address particularly well (or poorly) suited for that context?
3) What are the strengths and limitations of the outputs provided, and the implications for their use in policy analysis?
4) What are the most important research needs, in terms of methodological development, given your findings?

Disclaimer: The views expressed herein are those of the author(s) and do not necessarily reflect the views of the Economic Research Service or U.S. Department of Agriculture.

*With the contribution of Tine Häld, Arie Havelaar, and other members of the WHO Foodborne Disease Burden Epidemiology Reference Group (FERG) Source Attribution Task Force and of Willy Aspinall and Roger Cooke.
“A consistent and comparative description of the burden of diseases and injuries and the risk factors that cause them is an important input to health decision-making and planning processes. Information that is available on mortality and health in populations in all regions of the world is fragmentary and sometimes inconsistent. Thus, a framework for integrating, validating, analysing and disseminating such information is needed to assess the comparative importance of diseases, injuries and risk factors in causing premature death, loss of health and disability in different populations. Countries can combine this type of evidence along with information about policies and their costs to decide how to set their health agenda.” (WHO 2013b).

In 1990, recognizing that there was a need for internally consistent, comprehensive regional and global assessments of the physical impacts of major diseases and the risk factors contributing to their incidence, the World Health Organization (WHO) commissioned the first Global Burden of Disease (GBD) study (WHO 2013b). The study was updated in 2004 and again in 2010 (WHO 2013b). The GBD studies are large, worldwide efforts in data and research synthesis each involving hundreds of scientists. The 2010 GBD study involved 488 authors from 50 countries (Institute for Health Metrics and Evaluation (IHME) 2013).

The first GBD study in 1990 study developed estimates of the burden of mortality and morbidity by age, sex and region for 100 diseases and injuries for eight regions of the world. It introduced the use of the of the Disability-Adjusted Life Year (DALY) as an integrated, quantitative measure of the burden of each disease or injury type and attributed this burden to risk factors associated with each disease or injury type. A 2000-2002 update expanded on the analysis of risk factors, developing and applying a consistent analytic framework known as Comparative Risk Factor Assessment (CRA). These estimates were updated in 2004 and most recently in 2010. http://www.who.int/healthinfo/global_burden_disease/about/en/index.html. By measuring all health outcomes using a common metric, the Disability Adjusted Life Year (DALY), GBD estimates provide a means of comparing the relative public health burden of a diverse set of diseases around the world. This information makes it possible to assess the overall status of health around the world and to measure progress in improving health.
In 2006, the WHO convened an effort to assess the global burden of foodborne disease. This is one of the last major classes of disease to be added to the GBD initiative. A primary reason for the delay in estimating the GBD of foodborne disease has been the difficulty of identifying and estimating risk factors, in particular attributing disease to foodborne exposure and foodborne disease to consumption of particular foods. The Global Burden of Foodborne Disease study is a separate effort from the 2013 updating of the overall GBD estimates currently underway.

This paper focuses on the specific challenges of using expert elicitation to develop globally comparable risk attribution estimates. The paper begins by laying out data and research synthesis principles adopted in the GBD Study Protocol. It then briefly describes the specific data challenges faced in the Global Burden of Foodborne Disease Study (GBFD) and the ways research synthesis methods were used in the study as a whole. It then turns to the use of expert elicitation in the GBFD study; first looking at the criteria used to decide when to use expert elicitation rather than other research synthesis methods and then turning to the decisions that were made in developing a global expert elicitation. From this experience it draws insights into the four focus questions for this workshop:

5) What criteria should be used to evaluate the applicability of different research synthesis methods to particular types of problems and data?

6) What particular characteristics of the problem and data make the research synthesis method(s) you address particularly well (or poorly) suited for that context?

7) What are the strengths and limitations of the outputs provided, and the implications for their use in policy analysis?

8) What are the most important research needs, in terms of methodological development, given your findings?


Methodologically, the GBD project is perhaps most associated with the development and global application of DALY metrics. The DALY provides a non-monetary measure of the aggregate burden of radically different disease outcomes, e.g., mild diarrhea and death (Murray 1994). But underlying application of this aggregate health metric has been an immense effort at
synthesizing scientific data and research on health outcomes literally at a global scale. The health outcomes whose burden are measured by DALYs are estimated using more conventional modeling of disease incidence and risk factors contributing to this incidence around the world.

One of the primary methodological features of the GBD studies has been establishment of consistent principles to guide all work on GBD estimates and procedures for making decisions about changes in methodology. This provides a means of coordinating work across the very large number of collaborators required for an estimation and modeling effort of the scope of the GBD studies.

Box 1 presents underlying principles and methodological guidance for the most recent set of data synthesis principles from the GBD 2013 Protocol (IHME 2013). The GBD 2013 Protocol is specifically guidance for updating the 2010 GBD estimates, but it also reflects guidance from the GBD 2010 protocol, which informed methodology for the Global Burden of Foodborne Disease Study. Together these illustrate attention to the problem of using very diverse data sources, often collected using different definitions and data collection procedures to develop consistent estimates that are defensible from a statistical perspective. These sources might include administrative data on vital statistics, survey and census data collected under rigorous sampling or data collection protocols, passive surveillance data and active surveillance data. Results and data bases from previously published and unpublished studies are all viewed as part of the data environment on which researchers must rely in modeling global incidence and disease burden. Both systematic review and meta-regressions are specifically identified as playing central roles in this effort (IHME 2013). One research synthesis method that is not mentioned is expert elicitation.

2. Data Challenges in the Global Burden of Foodborne Disease Study

In 2006, the WHO convened an effort to assess the global burden of foodborne disease (Lake et al. 2013). As mentioned above, a primary reason for the delay in estimating the GBD of
foodborne disease has been the difficulty of identifying and estimating risk factors, e.g., attributing disease to foodborne exposure and foodborne disease to consumption of particular foods (Ezatti 2012). This risk factor analysis is critical to the usefulness of GBD estimates in public health policy. Estimating the association between food consumption and foodborne illness is a difficult task, even in high income countries with sophisticated disease surveillance systems. Many diseases that are foodborne can also be contracted through non-food routes of exposure. Disease onset can lag exposure to pathogens by days or even weeks and chemicals by months or years. Individual recall of food consumption is notoriously unreliable after as little as 24 hours. To date, the capacity to sample and test foods for contaminants is quite limited even in as wealthy a country as the U.S. One of the best sources of data linking food consumption and foodborne illness comes from outbreak investigations, but even in the U.S. outbreaks account for less than 10% of all foodborne illnesses due to pathogens. Further, there is evidence that sporadic cases do not always follow the same pattern of food associations as outbreak cases. In the U.S. and other high income countries, active surveillance programs with associated case-control studies are being used to explore this question. But case-control studies are generally quite targeted and cannot by themselves provide comprehensive sets of attribution estimates. Research on methods to reliably attribute specific foodborne diseases to relevant food sources has been emerging over the past decade (Häld et al 2004, Batz et al. 2005, Hoffmann et al. 2007, Havelaar et al. 2008, Evers et al. 2008, Mullner et al. 2009, Little et al. 2010, Batz et al. 2012, Painter et al 2013). Currently, a significant coordinated research effort is underway within the U.S. federal government.

The Foodborne Disease Epidemiology Reference Group is responsible for organizing the overall Global Burden of Foodborne Disease study to assure that the multiple pieces of the work feed into a comprehensive modeling effort (Fig. 3 from WHO FERG report). The Reference Group is organized into 4 task forces: 1) infectious disease; 2) chemicals and toxins; 3) source attribution; and 4) country burden and disease protocols.

The Attribution Task Force is responsible for producing estimates that: 1) attribute incidence of disease that could be foodborne to major exposure routes (e.g., food, water, animal contact, human contact, and soil); 2) attribute incidence of foodborne disease to specific foods; and 3)
estimate the likelihood that specific foods that were deemed routes of foodborne disease exposure were contaminated at various stages of production, processing and food preparation. Attribution of disease to sources of exposure is central to producing estimates of foodborne disease. For example, WHO (2008) estimates that diarrheal disease kills 2.2 million people worldwide in a typical year, but it is currently unknown how many of those deaths are due to food, water or other exposures. This knowledge is important to the effectiveness of efforts to prevent these deaths. To fill the modeling needs of the larger global burden of foodborne disease study, separate attribution estimates are needed for each of 6 WHO regions.

The Foodborne Disease Burden Epidemiology Reference Group (FERG) Attribution Task Force has responsibility for choosing methods to be used in developing source attribution estimates. One of the first actions taken by the Task Force was commissioning a review of food attribution research worldwide (Pires 2013). Pires (2013) reviewed the emerging research on alternative methods of attributing illness to food sources. The purpose of the review was to evaluate the data requirements of each approach and the implications of these requirements for use in producing global attribution estimates. Pires identified six basic approaches to source attribution (Fig. 1): methods based on molecular subtyping, comparative exposure assessment, systematic review and meta-analysis of epidemiological studies, analysis of outbreak investigations, risk assessment, and expert elicitation.

Subtyping approaches rely on genetic diversity within a species of pathogen to be able to match contamination in food to a specific primary production source, usually on farm. They have proved useful in settings where there is good sampling from important likely sources, particularly animals on farm and food. The focus and precision of this method can contribute significantly to control and eradication efforts. Not all pathogens have sufficient genetic diversity for the method to be useful. It has been applied to *Campylobacter* in several countries, but has not proven as successful as its application to *Salmonella*. For the purposes of the GBD study, a major limitation is the lack of suitable sampling data in most regions of the world.
Comparative exposure assessment is also a laboratory-based method that relies on sampling of foods, as well as non-food routes of exposure. It has been used successfully to explore transmission of *Campylobacter* in the Netherlands (Evers et al. 2008) and New Zealand (Lake et al. 2007). Pires (2012) concluded that even though availability and uncertainty around exposure estimates were a major limitation in applying the approach globally, the wider availability of this type of data, even in some developing countries, made this a more tractable approach for at least some pathogens transmitted through a limited number of food routes. Her review also suggested that worldwide, enough exposure data existed for some chemical hazards, including aflatoxin, cadmium, dioxins and lead, that the approach could be used to produce attribution estimates for the GBD study.

Standard epidemiological studies including case-control and cohort studies have been used to investigate the association between food-associated risk factors and specific hazards (Pires 2013). Pires (2013) found that these are particularly useful in understanding risk factors underlying sporadic illness; but that their usefulness in developing globally consistent, regionally-specific attribution estimates depended on having a sufficient number of studies in relevant regions to be able to conduct systematic reviews and, if possible, meta-analyses. Pires advised that sufficient numbers of such studies had been conducted for two pathogens, *Cryptosporidium spp.*, and *Giardia lamblia*, around the world to use systematic reviews to estimate four attributable fractions for these pathogens. She also noted that cohort studies of lead exposure could be useful in global burden of foodborne disease source attribution.

Attribution based on outbreak data has been used in many countries, particularly for bacteria, viruses, and some parasites. Outbreaks are defined as the occurrence of 2 or more illnesses linked to a single exposure source. All high income countries and many developing countries regularly investigate disease outbreaks with a focus on identifying the pathogen cause and source of exposure that caused the outbreak. By their nature outbreaks involve acute illness and therefore are generally not useful for identifying the sources of chronic disease particularly from chemical sources. Outbreak investigation data has been used successfully to estimate the fraction of outbreaks due to foodborne exposure and to identify the fraction of foodborne
outbreaks that were linked to specific foods. Pires (2013) suggested that aggregation of outbreak data regionally, could support source attribution estimates even where there are few outbreaks each year at a country level.

Pires (2013) identified expert elicitation generally as a means of overcoming data limitations. Yet, she provided little discussions of the challenges involved in conducting expert elicitations at a global, yet regionally specific level or the limitations inherent in the method.

3. Attribution Methodological Choices and Expert Elicitation

The FERG Attribution Task Force used a tiered approach to determining the method to use to provide attribution estimates for each of 100 biological and chemical hazards (WHO 2007). Eight systematic reviews have been published as of late summer 2013 to help determine the adequacy of existing research to support FERG modeling, including attribution estimation (WHO 2013c). In some cases, there was sufficient research of a variety of types to make the Task Force comfortable with characterizing a hazard as foodborne and with identifying the food vehicle that caused exposure. This research could be of a variety of types, e.g., basic research on the microbial ecology of organisms, epidemiological research on human illness and sampling of foods. In some cases, the strength of evidence was such that a decision could be reached on the foodborne attribution estimate and the food vehicle without use of meta-analysis.¹ In the case of 20 biological and chemical hazards, the Task Force determined that there was inadequate data or prior research on which to base regional attribution estimates. For these hazards, the Task Force followed the Pires (2013) recommendation that expert elicitation be used.

Within the structure of the global burden of disease studies, basic methodological choices are made by the study leadership and tasks forces (Fig. 2). This is necessary to assure that individual research effort feed appropriately into the overall modeling effort. The FERG Attribution Task Force determined that to provide most policy-relevant information on source attribution, estimates would be developed at the WHO sub-region level. These estimates could

¹The final paper will expand on the use of systematic reviews and meta–analysis in determining attribution percentages. The focus of this paper is on the use of expert elicitation.
be aggregated to the WHO regional level, but would also be more useful in health policy and food safety management decisions. In addition, the Task Force felt that the greater homogeneity of conditions in smaller regions would make it easier for experts to provide realistic attribution estimates.

The Attribution Task Force also decided that the study should use the Cooke method of aggregating expert judgments based on performance weights (Cooke and Goossen 2007). In the Cooke method, performance weights are based on responses to “calibration questions”. Calibration questions are questions about unknown, but knowable variable values in substantive areas related enough to the actual questions of interest that the study participants will feel they can respond. The purpose of the calibration questions is not so much to evaluate experts’ substantive knowledge, as to evaluate how accurately and informatively they can characterize their own subjective uncertainty about unknown values. Each expert on a panel is asked the same set of 8-15 calibration questions. Experts are asked to provide a “best” central judgment, defined as the median of the distribution and a 90% credible interval for each question. If an expert provides “accurate” probability assessments, the true value of half the calibration variables will be above the median values provided by the expert and half will be below; 90% will fall within the expert’s 90% credible intervals. Cooke’s method weights respondents’ performance based both on accuracy in this sense and “informativeness” in the sense of having narrower, though adequately wide, credible intervals (Cooke and Goossen 2007). Respondents who consistently provide 90% credible intervals that are too narrow to capture the true value of the calibration variable are penalized more heavily than respondents with wide, and therefore less “informative”, credible intervals, that nonetheless capture the true value of the calibration variable.

Ideally, in the Cooke method, calibration questions ask about quantitative variables whose values will only be known in the future (Fig. 3). For example, in a study of fish populations in the U.S. Great Lakes, the future number of landings of a particular fish species in Lake Superior was used as a calibration variable. At times this is not possible, and in these cases, Aspinall has used questions about previously published data, comparisons of previously published data and
results of published meta-analysis results. Calibration questions are administered in in-person interviews in part to control access to information. This is particularly important with retrospective rather than prospective calibration questions.

Cooke finds that panels of more than 9 to 10 experts add little to the accuracy and informativeness of estimates using his aggregation method (Cook and Goosens 2007). A basic decision in the design of the GBFD elicitation was how much heterogeneity could be captured by a single panel of experts. This determination turned on the variability of conditions influencing exposure from the perspective of the biological and physical nature of the hazard and the variability in environmental conditions. For example, it is expected that exposure to common enteric pathogens in developing countries is heavily influenced by water and sanitation. In contrast, exposure to the same enteric pathogens in high income countries might be more related to failures in food and water safety or hygiene during food preparation that is unrelated to the availability of good water and sanitation facilities. The study design question becomes how broad an understanding experts are likely to have regarding these underlying conditions in various regions and their influence on transmission of the pathogens of interest. This requires an understanding or investigation of the range of experience and knowledge of typical experts. It also requires a basic understanding of heterogeneity in geographic conditions and their likely influence on the variables of interest.

Table 1 shows the decisions made in addressing this issue. These decisions reflect two primary considerations. First, what is the appropriate geographic scale for eliciting variables of interest. Second, what is the structure of the scientific profession in the sense of how is geographical and subject matter expertise organized.

There was considerable discussion of the appropriate sub-region structure to use for the elicitation. Different available international regionalization schemes reflect different factors that are likely to influence the sources of foodborne illness. A choice had to be made among these. Food consumption patterns clearly are a factor in determining exposure routes. The FAO Global Environment Monitoring System (GEMS) system regions, originally developed to predict radionuclides exposure through foods following the Chernobyl accident. However, water and
sanitation, food production and processing practices can be important determinants of contamination of foods. GEMS regions do not reflect the economic and institutional factors that affect the capacity to develop strong preventive systems (Fig. 4). The alternative chosen was the WHO sub-regions based on child mortality (Fig. 5). The thinking behind this choice was that factors influencing child mortality include overall levels of economic development, political/institutional development, and other factors such as general educational attainment that could affect the capacity to maintain hygiene and food handling systems. Another advantage of the WHO sub-regions is that they tend to be geographically contiguous, while the most recent GEMS regions combine countries from multiple continents.

Initial discussions assumed that scientific expertise and professional knowledge was also tied to region in the sense that those who live in a region will necessarily be most expert in that region. After some investigation, it became apparent that domicile was not the only factor affecting regional expertise. For example, many recognized experts on sanitation and on food safety in developing countries were based in developed countries. A focus that tied expertise to domicile would have missed this link. This discussion raised the further question as to whether there were areas in which the hazard was more critical than regional environmental and institutional variation in determining appropriate expertise for the elicitation. The initial structure of the panels was by region. Once these possibilities were raised, there was a realization that for many hazards it would be possible to have a single global panel that could provide estimates for all sub-regions. In the end, it was decided that regionally specific panels would only be used for a group of common enteric pathogens (Table 1).

Another factor influencing the regional structure of the panels was respondent burden. Given that judgments are needed for 9 “other” enteric pathogens in 11 sub-regions in the developing world, the burden was seen as too great for a single panel of 9 to 10 experts (Table 1). In order to protect the quality of judgments, we are currently considering a variety of options for reducing burden. These include reducing the number of sub-regions for which a single panel must give responses or reducing the number of pathogens for which experts must give judgments. The most likely outcome is that FERG will decide to reduce the number of sub-
regions. For example, rather than a single panel providing judgments for all 11 developing country sub-regions, there may be a panel that gives judgments for the WHO sub-regions in Africa and the Eastern Mediterranean. The ability to do this is dependent on the number of experts who are willing to participate in the elicitation.

4. Lessons for Future Research

This workshop raises a series of questions about the more generalizable lessons that can be drawn from particular applications for research synthesis. Both the more general work of the FERG Attribution Task Force and the specific experience of developing a globally consistent, regional set of attribution estimates based on expert judgment address these questions.

1. What criteria should be used to evaluate the applicability of different research synthesis methods to particular types of problems and data?

This research effort illustrates some established criteria for when to use expert elicitation as well as some that are not usually stated and may be uniquely illustrated by the problem of food source attribution. As in general, expert elicitation was turned to in this GBD study only when other, more conventional options, including systematic reviews and meta-analysis were unavailable. If there is secondary data on which to base estimates, if original data collection is feasible, if there is an adequate body of existing research on which estimates can be developed through systematic reviews and meta-analysis, then any of these options should be used before turning to expert elicitation.

The problem of food source attribution shows that there are cases not only where data is not currently available, but also where it may be very difficult to develop research designs that address critical research questions. These situations are candidates for expert elicitation. The other, often unstated, requirement is that there is knowledge that experts can draw on to provide the needed estimates. In particular, the problem of food source attribution illustrates that expert elicitation is particularly well suited to situations where relationships of interest are affected by a complex set of conditions which interrelate and about which experts has a basis for reasoning. So in the case of food attribution, multiple factors all combine to influence the
likelihood that an illness was foodborne or that a particular food was the source of exposure. These include the immunological health of the population, food consumption patterns, and factors that influence the likelihood that food is contaminated, such as food production and processing patterns, water and sanitation, and environmental conditions. Risk assessment models this kind of information synthesis formally, but experienced experts can informally do a similar synthesis exercise.

2. What particular characteristics of the problem and data make the research method you address particularly well suited for that context?

In the case of the GB of foodborne disease study, it was not practicable to develop formal risk assessment modeling to produce the number of attribution estimates needed for modeling risk factors associated with the burden of foodborne disease. In the cases where expert elicitation was used, data that would have been needed for such modeling was lacking. Further, it would have been difficult to develop modeling that would produce comparable estimates across regions given the variation in available data as well as the considerable variation in real world conditions. The need to produce attribution estimates within a meaningful time line was also a factor.

3. What are the strengths and limitations of the outputs provided and the implications for policy analysis?

It follows from the above discussion that a major strength of expert elicitation is the ability to provide a consistent set of estimates across data environments and real world conditions that vary considerably. A limitation is that expert elicitation is not empirical data or empirical estimates. It would be preferable to be able to actually measure the relationship between food and illness, but this is difficult. Because of this, winning the trust of the scientific community in the use of expert elicitation is a challenge with the method. The use of formal, transparent methods that draw on best practices from related fields, like survey design and research on risk perception and evaluation helps as well as focusing on improving and evaluating expert elicitation methodology will all help in making the case for the method. It also helps to remind
the scientific audience that expert elicitation is not a substitute for data, but can help focus and guide future research and enhance the effectiveness of data collection efforts.

4. What Are the Most Important Research Needs, in Terms of Methodological Development, Given Your Findings?
Experience with this effort at using expert elicitation on a very large scale provided a number of insights into research needs. First, there is a need for more research on judgment aggregation methods. Second, if expert elicitation is to be conducted in large numbers or at a large scale, then methods other than in-person facilitation will be needed. Solid research on the influence of in-person versus self-administered approaches on judgments, uncertainty, and willingness to participate in the study are needed to determine best practices. Similarly, there is a need to development of additional approaches to administering expert elicitations without in-person facilitation. Third, with the Cooke method, there is a need for research comparing the relative usefulness of prospective and retrospective calibration questions in evaluating probability assessors’ accuracy and informativeness. There is also need for research into the how closely related calibration questions need to be to the variables of actual interest for the calibration questions to be useful in evaluating experts ability to assess uncertainty in their probability judgments.
Box 1. 2013 Global Burden of Disease Study Methodology Guidelines

GBD 2013 Selected Key Principles and Assumptions

Uncertainty
1. Because the GBD 2013 produces estimates for a mutually exclusive and collectively exhaustive set of disease and injury causes, it is important to convey to users the strength of the evidence for each quantity through the reporting of uncertainty intervals.
2. The GBD 2013 estimates uncertainty distributions for each quantity and reports and visualizes various metrics of uncertainty, such as 95% uncertainty intervals.

Internal Consistency
1. The sum of cause-specific mortality must equal all-cause mortality following ICD underlying cause rules.
2. The sum of cause-specific estimates of impairments, such as blindness, must equal estimates of all-cause impairments.
3. Where we believe incidence, prevalence, remission, duration, and excess mortality are not changing over time we require rates to be internally consistent

Iterative Approach to Estimation
1. New data and methodological innovation will lead to revision of estimates.
2. Burden of disease estimation is an iterative process. Revisions will result in a re-estimation of the entire time series so that results are always available over time using consistent data and methods.

Data Synthesis Principles
1. We will identify all available relevant sources of data for a given disease, injury, and risk factor and for all-cause mortality.
2. For all data sources identified, we will assess the sampling method, case definitions, and potential for bias.
3. For cause of death data, we will map variants of the ICD and will redistribute garbage codes.
4. For data on incidence, prevalence, remission and excess mortality, we will use statistical methods to characterize the relationship between different case definitions, diagnostic technologies, recall periods, etc.
5. We will use these relationships to transform data into comparable units, definitions, or categories. Wherever possible, we will propagate uncertainty in these mappings into the uncertainty interval for the measurement.
6. Some measurements may have to be excluded because they cannot be made comparable to the rest of the measurements or have fundamental problems of validity.
7. We will synthesize all the appropriate data using statistical methods that can handle both sampling and non-sampling error.
8. The statistical methods employed will improve predictions where data are sparse by allowing for use of covariates and by borrowing strength across time or geography.
9. All estimates will be generated with 1000 (or more) draws of the quantity of interest from the posterior distribution.
10. Where possible, we will demonstrate validity of the statistical methods by using out-of-sample prediction.
(IHME 2013 pp. 6-7).
Section 5. Estimation Flow

Ongoing GBD estimation will follow the methodology presented in the GBD Study 2010^{4-8}, unless otherwise approved by the GBD Scientific Council. The following flowchart illustrates the flow of the key components of the GBD estimation process. The text below the flowchart provides additional detail and is numbered to match the numbered component(s) of the flowchart it describes.

5-7. Cause of death database
The GBD 2013 will use the expanding cause of death database maintained by IHME, which maps data across various revisions and national variants of the International Classification of Diseases and Related Health Problems (ICD). Incorporated into this database are all relevant sources of cause of death data, including vital registration, verbal autopsies, census and survey data, police records, hospital data, surveillance systems, and population based registries for specific diseases. (emphasis added)

10. Disease sequelae epidemiology data
To support estimation for each disease sequelae incidence, prevalence, duration, remission, and excess mortality, a database of available published and unpublished data will be developed and maintained by IHME. Key inputs into this database include systematic reviews of the published and unpublished literature, analysis of household survey data, antenatal clinic surveillance data, reportable disease notifications, disease registries, hospital admissions data, outpatient visit data, population-based cancer registries, active screening data, and other administrative data.
11. Estimating disease sequelae prevalence, incidence, and duration
Most estimates for disease incidence, prevalence, duration, and excess mortality will be calculated using the latest version of the *GBD Bayesian meta-regression tool* (DisMod-MR or subsequently approved version), which was designed to address some of the key challenges in burden of disease analysis. Core Analytic Team members will carry out this modeling. For some causes where more complicated models capturing more stages of disease progression are necessary, such as HIV, more elaborate natural history models will be used, refined, or developed.

18a-c. [19]. Risk factor exposure database
Analogous to the same principles used for work on disease sequelae, IHME has created and will continue to expand a *database of published and unpublished sources on the prevalence of exposure*. For some risks, innovative sources such as satellite imagery will be used following the experience in GBD 2010. For a number of risk factors, primary survey data will be collated and re-analyzed along with published studies.

19. Estimating the prevalence of exposure
Exposure estimates will be developed for many risks using the *GBD Bayesian meta-regression tool* DisMod-MR, or a subsequently developed and approved version. For some risks such as ambient air pollution, alternative modeling strategies will be used. In all cases, the estimation of exposure prevalence will generate uncertainty distributions. (IMHE 2013 pp. 24-28, *emphasis added*).
Figure 1. Principles and Data Requirements for Food Source Attribution Methods (Pires 2013)

<table>
<thead>
<tr>
<th>Source attribution method</th>
<th>Principle</th>
<th>Data requirements</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Occurrence approaches</td>
<td>Compare the subtypes of isolates from different sources (e.g., animals, foods) with the same subtypes isolated from humans.</td>
<td>Characterization of the hazard by subtyping methods (e.g., phenotypic or genotypic subtyping). Collection of temporally and spatially related isolates from humans and various sources.</td>
<td>Salmonella: Held et al., 2004; Pires and Hold, 2010; Campylobacter: Mullner et al., 2009a; Wilson et al., 2008</td>
</tr>
<tr>
<td>Comparative exposure assessment</td>
<td>Determine the relative importance of the known transmission routes by estimating the human exposure to the hazard via each route.</td>
<td>Occurrence of the hazard (prevalence and concentration) in all putative sources; information on the changes of the level of the hazard in the main steps of the transmission chain; human exposure data, typically food consumption data.</td>
<td>Campylobacter: Evers et al., 2008; Listeria: FDA, 2003; Lead: EFSA, 2010</td>
</tr>
<tr>
<td>Epidemiological approaches Analysis of data from sporadic cases*</td>
<td>Case-patients are interviewed and asked about exposures. Case-control studies (CCS): case-patients and a representative group of asymptomatic individuals are interviewed, and the relative role of exposures is estimated by comparing the frequency of exposures among cases and controls. A systematic review (SR) and meta-analysis of CCS combines information from all available studies; useful for regional studies.</td>
<td>Register data on human cases is available. SR of epidemiological studies: sufficient number of studies focusing on the same hazard conducted and made publically available. Studies can originate from multiple regions and time periods.</td>
<td>E. coli O157: Voetsch et al., 2007; Salmonella/ Campylobacter: Domingoae et al., 2012a, 2012b</td>
</tr>
<tr>
<td>Analysis of data from outbreak investigations</td>
<td>Attribute illnesses to food sources on the basis of the number of outbreaks that were caused by each food.</td>
<td>For each outbreak reported, information on the date, number of suspected and confirmed cases and implicated source. Number of human cases of disease before and after intervention/natural change.</td>
<td>Multiple pathogens: Batz et al., 2012; Ravel et al., 2009; Salmonella: Pires et al., 2010b; E. coli: Gupta et al., 2007; Campylobacter: Vellinga and Van Loooc, 2002</td>
</tr>
<tr>
<td>Intervention studies and natural experiments</td>
<td>Evidence of the burden of a foodborne disease attributable to specific sources by measurement of the impact of an intervention in a specific source facilitated if the intervention study has a randomized design and is carried out in a controlled environment. Natural experiments: a change in exposure or consumption behavior of the population results in a decrease in the number of reported cases.</td>
<td>NA</td>
<td>Multiple pathogens: Havelaar et al., 2008; Ravel et al., 2010; Salmonella: Hoffmann et al., 2007</td>
</tr>
<tr>
<td>Expert elicitation</td>
<td>Answers of experts selected based on defined criteria and obtained through structure protocols are analyzed statistically.</td>
<td>NA</td>
<td></td>
</tr>
</tbody>
</table>

*Includes case-control studies, case-series studies, and cohort studies. NA, nonapplicable.
Table 2. **WHO Food Source Attribution Expert Elicitation Study Panels by Hazards**

<table>
<thead>
<tr>
<th>Task Force</th>
<th>Panels</th>
<th>Hazards (# and list)</th>
<th># of Expert Panels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#</td>
<td>List</td>
<td></td>
</tr>
<tr>
<td>Chemical</td>
<td>Metals</td>
<td>3 Global: Inorganic arsenic, lead, cadmium</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Dioxin</td>
<td>1 Global:</td>
<td>1</td>
</tr>
<tr>
<td>Parasites</td>
<td>Protozoa</td>
<td>3 Global: Crypto, Entamoeba histolytica, Giardia</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Echinococci</td>
<td>2 Global: E. granulosus, E. multilocularis,</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Toxoplasma</td>
<td>1 Global: Toxoplasma gondii,</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Ascaris</td>
<td>1 Global: Ascaris lumbricoides</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Brucella</td>
<td>1 Global panel</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>HEV A</td>
<td>1 Global panel</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>Developed: Other diarrheal</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>Developing: Other diarrheal</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>12</td>
</tr>
</tbody>
</table>


Fig. 2  Structure of the WHO Foodborne Disease Burden Epidemiology Reference Group (FERG) (WHO 2007).

Figure 3. Example of a Prospective Calibration question.

**Background:** The New Zealand Ministry of Health publishes annual reports on surveillance of notifiable diseases in New Zealand. New reports are released in April. The most recent data is for 2012 cases. Source: ESR, New Zealand Public Health Observatory, Notifiable Diseases, http://www.nzpho.org.nz/NotifiableDisease.aspx.

4.3.2. What was the change in the rate number of human campylobacteriosis cases per 100,000 population in New Zealand from 2006 to 2013?

*Your response*

<table>
<thead>
<tr>
<th>Percentile</th>
<th>low (5th)</th>
<th>best (50th)</th>
<th>high (95th)</th>
</tr>
</thead>
</table>
Figure 4. FAO Global Environment Monitoring System (GEMS) Food Cluster Diet Regions 2012
Figure 5. WHO Subregional Country Groupings for the Global Assessment of Disease Burden

Figure 6. Example Attribution Exercise for Exposure Pathway Attribution.

Table 1: Pathogen X Total Exposure
Percent of All Human Cases in a Typical Year

<table>
<thead>
<tr>
<th>Point of Exposure*</th>
<th>lower credible value (5th percentile)</th>
<th>central value (50th percentile)</th>
<th>upper credible value (95th percentile)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food</td>
<td>10%</td>
<td>20%</td>
<td>30%</td>
</tr>
<tr>
<td>Animal Contact</td>
<td>0%</td>
<td>5%</td>
<td>12%</td>
</tr>
<tr>
<td>Human to Human Contact</td>
<td>5%</td>
<td>10%</td>
<td>28%</td>
</tr>
<tr>
<td>Water</td>
<td>50%</td>
<td>65%</td>
<td>80%</td>
</tr>
<tr>
<td>Other</td>
<td>0%</td>
<td>2%</td>
<td>5%</td>
</tr>
</tbody>
</table>

*Think of the source that was the direct cause of human exposure. We are not asking about how the source was contaminated in this table.

Figure 7. Example Attribution Exercise for Food Source Attribution

Table 2: Pathogen X Foodborne Exposure
Percent of Foodborne Cases in a Typical Year

<table>
<thead>
<tr>
<th>Food Consumed**</th>
<th>lower credible value (5th percentile)</th>
<th>central value (50th percentile)</th>
<th>upper credible value (95th percentile)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beef</td>
<td>10%</td>
<td>15%</td>
<td>25%</td>
</tr>
<tr>
<td>Goat, lamb and other small ruminants’ meat</td>
<td>3%</td>
<td>5%</td>
<td>18%</td>
</tr>
<tr>
<td>Dairy (milk and milk products)</td>
<td>0%</td>
<td>5%</td>
<td>8%</td>
</tr>
<tr>
<td>Pork</td>
<td>25%</td>
<td>35%</td>
<td>50%</td>
</tr>
<tr>
<td>Poultry Meat</td>
<td>9%</td>
<td>10%</td>
<td>12%</td>
</tr>
<tr>
<td>Vegetables (excluding dried legumes)</td>
<td>4%</td>
<td>15%</td>
<td>20%</td>
</tr>
<tr>
<td>Fruits</td>
<td>1%</td>
<td>5%</td>
<td>10%</td>
</tr>
<tr>
<td>Nuts</td>
<td>0%</td>
<td>1%</td>
<td>2%</td>
</tr>
<tr>
<td>Other foods</td>
<td>1%</td>
<td>5%</td>
<td>10%</td>
</tr>
</tbody>
</table>

**Attribute cases of illness to the foods that were already contaminated when they entered the home kitchen or other place of final food preparation. Do not consider cross-contamination in the home kitchen or other place of final food preparation.
References


Ezatti, M., personal communication with author, fall 2012.


World Health Organization. 2013b. About the Global Burden of Disease (GBD) project. 

World Health Organization. 2013c. Initiative to estimate the Global Burden of Foodborne Diseases: 
Information and Publications. 

Watts C, Cairncross S (2013). "Should the GBD risk factor rankings be used to guide policy?". 