 DOI: 10.1111/risa.12385

Perspective

Research Synthesis Methods in an Age of Globalized Risks: Lessons from the Global Burden of Foodborne Disease Expert Elicitation

World Health Organization, Foodborne Epidemiology Reference Group, Source Attribution Task Force*

We live in an age that increasingly calls for national or regional management of global risks. This article discusses the contributions that expert elicitation can bring to efforts to manage global risks and identifies challenges faced in conducting expert elicitation at this scale. In doing so it draws on lessons learned from conducting an expert elicitation as part of the World Health Organizations (WHO) initiative to estimate the global burden of foodborne disease; a study commissioned by the Foodborne Disease Epidemiology Reference Group (FERG). Expert elicitation is designed to fill gaps in data and research using structured, transparent methods. Such gaps are a significant challenge for global risk modeling. Experience with the WHO FERG expert elicitation shows that it is feasible to conduct an expert elicitation at a global scale, but that challenges do arise, including: defining an informative, yet feasible geographical structure for the elicitation; defining what constitutes expertise in a global setting; structuring international, multidisciplinary expert panels; and managing demands on experts' time in the elicitation. This article was written as part of a workshop, "Methods for Research Synthesis: A Cross-Disciplinary Approach" held at the Harvard Center for Risk Analysis on October 13, 2013.

KEY WORDS: Disease burden; expert elicitation; expert judgment; exposure estimates; foodborne illness; research synthesis; source attribution; systematic review; uncertainty quantification

1. INTRODUCTION

We live in an age that increasingly calls for management of risk at a global scale. Increased global travel and trade heighten the likelihood that risks, such as the 2014 Ebola outbreak, can spread globally. Growing scientific understanding of how natural and technological systems function at a global scale has increased awareness of risks, such as climate change and mercury pollution, that must also be managed at

Chair: Tine Hald (National Food Institute, Technical University of Denmark). SATF and FERG Task Force chairs: Fred Angulo (U.S. Centers for Disease Control and Prevention), Wan Mansor Bin Hamzah (Disease Control Division, Malaysian Ministry of Health), David Bellinger (Harvard School of Public Health), Robert Black (The John Hopkins University Bloomberg School of Public Health), Nilanthi de Silva (Faculty of Medicine, University of Kelaniya), Dörte Döpfer (School of Veterinary Medicine, University of Wisconsin– Madison), Arie Havelaar (Emerging Pathogens Institute, University of Florida), Herman Gibb (Gibb Epidemiology Consulting LLC), Fumiko Kasuga (National Institute of Health Sciences, Japanese Ministry of Health, Labour and Welfare), Rob Lake (Institute of Environmental Science and Research (ESR) Ltd.), Muhammad B Rokni (Teheran University of Medical Sciences), Niko Speybroeck (Université catholique de Louvain). Commissioned scientists/resource advisers/consultants: Willy Aspinall (Bristol University and Aspinall & Associates), Roger Cooke (Resources for the Future), Sandra Hoffmann (U.S. Department of Agriculture, Economic Research Service), Brecht Devleesschauwer (Emerging Pathogens Institute, University of Florida), Sara M. Pires (National Food Institute, Technical University of Denmark).

*Address correspondence to Sandra Hoffmann, U.S. Department of Agriculture, Economic Research Service, Washington, DC, USA; shoffmann@ers.usda.gov.

 0272-4332/16/0100-0001$22.00/1

a global scale. Even when risks themselves are not transmitted globally, their impacts may have global repercussions. International efforts to reduce local outcomes that have global impacts, like malnutrition in developing countries, need to be informed by globally comparable data.

Management of risks at a global scale requires coordinated action by actors at the local, national, or regional levels and therefore also requires information that is meaningful at the local, national, or regional level, and comparable globally. Yet disparities in economic and institutional development around the world result in great geographic differences in the availability, quality, and form of data and scientific research. Research synthesis methods, particularly expert elicitation, can provide a means of bridging these gaps in data and scientific literature, providing a sounder informational basis for many global risk management decisions.

On October 13, 2013, the Harvard Center for Risk Analysis held a workshop, "Methods for Research Synthesis: A Cross-Disciplinary Approach." The workshop's goal was to gain a better understanding of the role research synthesis methods play in risk analysis and the relative strengths and limitations of different methods. Robinson and Hammitt[1] draw out lessons from the workshop and subsequent publications, including this article and a special issue of *Risk Analysis* in June, for future developments and applications of research synthesis methods.

This article's primary objective is to identify potential contributions that expert elicitation can bring to characterizing global risks and to identify challenges in conducting expert elicitation at a global scale. It does so by examining a new effort to use expert elicitation as part of the World Health Organization's (WHO) initiative to develop the first estimates of the global burden of foodborne diseases; a study commissioned by the Foodborne Disease Epidemiology Reference Group (FERG).[2,3] As discussed below, prevention of foodborne disease is increasingly a global risk management challenge. The WHO estimates are needed to inform both international and national management of foodborne diseases. These estimates will be provided for each WHO region and will be globally comparable. Other research synthesis methods in addition to expert elicitation, particularly systematic reviews, were used throughout the WHO global burden of foodborne disease initiative. In examining lessons learned about expert elicitation to inform management of global risks, this article will also look at how decisions were reached about when to use expert elicitation rather than other research synthesis methods.

2. OVERVIEW

2.1. Why Estimate the Global Burden of Foodborne Disease?

Foodborne disease is a significant cause of illness and death worldwide. Each year, 2.2 million people, 1.9 million of whom are children, die from diarrheal illness. Foodborne diseases are believed to account for a substantial proportion of these deaths.[4] While most of these deaths occur in developing countries, foodborne disease is also an important public health concern in high-income countries. Approximately 48 million out of a population of 310 million Americans have foodborne illnesses in a typical year; of these, over 3,000 die.[5,6] In the United States, the 14 leading foodborne pathogens alone impose over $14 billion in economic burden each year.[7]

The full extent of the burden of foodborne disease globally is currently unknown. Surveillance of foodborne disease is difficult, even in developed countries. For example, the U.S. Centers for Disease Control and Prevention (CDC) estimates that in the United States only, 1 in 29 illnesses caused by nontyphoidal *Salmonella spp.* are diagnosed and reported.[5] Data on the incidence in developing countries are even scarcer.[8]

Foodborne disease has become a global risk management problem. With increasing globalization in food trade, one country's food safety problems can quickly spread to others. Recent outbreaks of shigellosis in Australia and Denmark were caused by baby corn imported from Thailand.[2] Imported frozen dumplings from China caused organophosphate poisoning in over 3,000 Japanese consumers.[2]

Foodborne disease is also a global development issue. WHO maintains that substantial investment in foodborne disease prevention and control in developing countries is essential to making progress on development concerns underlying the U.N. Millennium Development Goals.[9] Foodborne disease perpetuates a cycle of poverty in developing countries by reducing labor productivity, increasing child mortality, and impairing efforts to improve nutrition (WHO 2008). Foodborne disease in developing countries also hampers economic development through impacts on tourism and food export industries.[10,11]

Consistent global estimates of foodborne disease are needed to help guide policy and prioritize public health and development investments and will provide a baseline against which the impacts of future interventions can be evaluated. International trade law and public health agreements require that countries base their food safety measures on sound science and risk analysis.(12) WHO has identified the lack of accurate data on the extent and cost of foodborne disease as a major obstacle to policymakers around the world setting "appropriate, evidence-based priorities in the area of food safety."(2,4)

2.2. What the WHO is Doing

The WHO launched the "Initiative to Estimate the Global Burden of Foodborne Diseases" to meet this need for globally consistent, regionally specific, information on foodborne disease.(2,4) WHO established a large multidisciplinary body of scientists from around the world, the Foodborne Disease Burden Epidemiology Reference Group (FERG), to implement this new initiative.(8,9) The FERG is tasked with developing global estimates of the foodborne disease burden for 40 major microbial, parasitic, and chemical causes of foodborne disease, stratified by sex, age, and WHO region. Burden is measured using the summary health metric, disability adjusted life years (DALY).(13) The DALY allows comparison of burden across widely differing diseases and can be used in cost-effectiveness analysis.(14,15) The FERG is also tasked with providing estimates that will allow foodborne disease burden to be attributed to food sources. Country case studies are being developed to help examine how individual countries can make use of these regional estimates in developing national estimates and in managing foodborne hazards at the national level.

The FERG is carrying out its efforts to estimate the global burden of foodborne disease through six task forces: (1) enteric diseases; (2) parasitic diseases; (3) chemical and toxin-caused diseases; (4) source attribution; (5) country studies; and (6) a computational task force. The three disease-specific task forces are responsible for assessing and synthesizing existing data on the burden of foodborne disease and providing global burden of disease estimates. Systematic reviews are being used extensively by the disease tasks forces to estimate the burden of disease.(16–19) The Source Attribution Task Force (SATF) is charged with estimating the foodborne proportion of disease burden imposed by each hazard and the contribution of specific foods to foodborne exposure; that is, providing source attribution estimates. The Country Study Task Force is developing protocols and other guidance for countries to use in developing their own national burden of disease studies. The Computational Task Force is integrating information from the other tasks forces into a consistent model and calculating food-specific burden of foodborne disease estimates for each WHO region.(20) Major methodological decisions are reached by consensus among FERG members in consultation with study leaders.(10) The remaining discussion focuses on the work of the SATF.

2.3. Defining the Source Attribution and Choosing Among Research Synthesis Methods

The SATF was charged with: (1) defining source attribution; (2) assessing currently available methods for attributing specific foodborne diseases, for example, salmonellosis, to sources of exposure; (3) proposing suitable source attribution methods for different causal hazards; and (4) developing primary exposure and food exposure source attribution estimates. The SATF defined source attribution differently for each of the two major types of attribution it was asked to provide: attribution to primary exposure pathways and to food exposure pathways (Table I). Primary exposure pathways for biological hazards were constructed to distinguish between exposure via food, water, soil, air, animal or human contact, and "other" primary exposure routes. For these primary pathways, the task force defined the exposure source as the pathway that was the direct cause of human exposure. To attribute foodborne disease caused by biological hazards to foods, the task force partitioned foods into 14 mutually exclusive categories (Table I). For food exposure pathways, the task force defined the source of exposure as the food that was contaminated with the disease-causing hazard at the time the food entered the place of final food preparation. This definition of food attribution sets aside the contribution of cross-contamination among foods during food preparation and focuses the policy use of the FERG burden estimates on prepreparation prevention of foodborne disease.

Scientists have many decades of experience modeling the role of foods in chemical exposures. Compared to methods used for estimating foodborne pathogen exposures, methods for modeling foodborne chemical exposure are well developed.(21) In contrast, estimation of the contribution of foods as

Table I. Possible Exposure Routes for Primary Pathways and Foodborne Exposure Pathways in the Global Burden of Foodborne Diseases Estimates (Biological Hazards)

Primary Exposure Pathways	Foodborne Exposure Pathways
Question addressed What was the primary exposure route?	Question addressed Which foods caused foodborne exposure?
Possible primary exposure routes	Possible foodborne exposure routes
Food	Beef
Animal contact	Ruminant meat
Human-to-human contact	Dairy
Water	Pork
Air	Poultry meat
Soil	Eggs
Other	Vegetables
	Fruits and nuts
	Grains and beans
	Oils and sugar
	Finfish
	Shellfish
	Seaweed
	Other foods

pathogen exposure pathways is a relatively new and rapidly developing area of science.(22,23) The SATF commissioned a review of research on foodborne pathogen source attribution.(22) This review identified five methods being used to estimates source attribution based on analysis of primary data: (1) comparison of subtypes of pathogens in humans and foods or food sources; (2) risk modeling/comparative exposure assessment; (3) epidemiological studies of sporadic cases; (4) analysis of data from outbreak investigations; and (5) intervention studies and natural experiments. When data are incomplete, the review found expert elicitation useful for attributing disease to primary exposure pathways and specific food exposures.(22)

The task force concluded that for many hazards data limitations prevented development of globally consistent, regional source attribution estimates based on primary data analysis. The task force also concluded that the depth of the existing literature was inadequate for systematic reviews to provide consistent estimates across WHO regions. The SATF treated some hazards as 100% foodborne (e.g., *Listeria monocytogenes*, *Mycobacterium bovis*, all foodborne trematodes, *Taenia solium*, *Trichinella* spp., cyanide in cassava and peanut allergens) based on the weight of the scientific literature, but not necessarily systematic reviews. For other hazards (e.g., aflatoxin, inorganic arsenic, cadmium, dioxin, and methyl mercury), adequate data on foodborne exposure existed to develop a risk assessment model to estimate foodborne exposure. Together, these two sets of decisions obviated the need to use expert elicitation to provide attribution estimates for 21 hazards. For the remaining 19 hazards, the SATF decided expert elicitation was needed to obtain source attribution estimates.

Practical and conceptual considerations influenced the SATF's choice among expert elicitation methods. Methods requiring group meetings or consensus were not used due to methodological concern about the influence of group dynamics on consensus, WHO's need for documentable transparency, lack of travel budget, and technological and time zone difficulties in conducting effective group meetings remotely across many regions of the world.(24) The task force chose Cooke's classical model as the method for the source attribution expert elicitation because of its successful use in prior food source attribution studies and its use of transparent, quantitative performance measures.(25–27) Cooke's classical model elicits judgments independently from individual experts and aggregates them on the basis of performance on calibration questions. The method uses performance weights and an aggregation method designed to maximize the accuracy and informativeness of aggregate estimates.(27,28)

2.4. Design of the WHO FERG Source Attribution Expert Elicitation

Remaining decisions about the design of expert elicitation were heavily affected by the global scope of the WHO FERG source attribution study. These decisions included: (1) choice of regional structure; (2) structuring of expert panels; and (3) development of calibration questions. Several themes recurred in making these decisions, including: the need for experts from multiple disciplines; the need to reduce cognitive challenge and manage demands on respondents' time; the need to represent the geographic heterogeneity of regions as accurately as possible; and the challenges posed by limited data and research. While these themes are common to many expert elicitation modeling efforts, the global scope of this study increases the challenge of addressing them. While other expert elicitations have involved experts from around the world, we know of none that have been used to provide regionally specific and globally consistent estimates as are needed by WHO.

2.4.1. Choice of Regional Structure

One of the most fundamental decisions in the study was how to structure the regions used in the elicitation. It is more cognitively challenging for experts to provide source attribution estimates for large heterogeneous regions than for smaller, more homogeneous regions. The WHO FERG initiative will provide disease burden estimates for each of six WHO regions. But WHO regions are large and heterogeneous in terms of factors that affect foodborne hazard exposures (Fig. 1a). For example, the WHO Region of the America's includes North America, the Caribbean, Central America, and South America, areas that differ significantly in terms of factors influencing exposure to foodborne hazards, such as access to improved water and sanitation, food processing conditions, and food transportation, storage, and handling conditions.

An alternative to eliciting estimates directly for each WHO region is to elicit them for smaller, more homogeneous regions and then aggregate these attribution estimates to the WHO regional level. Both Globally Environmental Monitoring System (GEMS) Food Consumption Custer Diet regions and WHO subregions were considered (Figs. 1b and c). Ultimately, a decision was made to elicit attribution estimates for regions developed for prior global burden of disease (GBD) studies.[30] With two exceptions, these 14 GBD regions are organized as subregions of the current WHO regions, though they are not official WHO subregions. For simplicity we refer to them as GBD subregions. These subregions subdivide WHO regions based on under-5 child mortality and adult mortality rates. Mortality rates, particularly under-5 mortality, are correlated with factors such as access to improved water and sanitation and basic development indicators that also correlate with factors influencing foodborne contamination and illness. GBD subregions are more geographically contiguous than GEMs regions and tend to be more homogeneous in terms of water and sanitation conditions, general development levels, and even basic dietary patterns than WHO regions (Figs. 1a and 1c). One-on-one interviews used in designing the elicitation instrument suggested that experts found it easier to think about source attribution for GBD subregions than for GEMS or WHO regions.

2.4.2. Structure of Expert Panels

Cooke's classical model collects individual judgments from panels of multiple experts and then aggregates individuals' judgments using performance weights.[27,28] A choice had to be made whether to structure expert panels by subregions with multiple hazards or by hazards for multiple subregions. Scientific expertise related to foodborne disease and food safety is usually developed around hazards. Many subject matter experts work in multiple subregions, often including subregions other than those in which their home institution is based. As a result, panels were organized by hazard and experts were asked to give judgments for multiple subregions.

An exception to the use of global panels was made for diarrheal enteric pathogens (Table II). Experts on enteric pathogens tend to have professional experience with multiple pathogens, but that expertise tends to be more regionally specialized, particularly between countries with high and low child mortality rates, than is the case for the nonenteric hazards included by the WHO FERG. The panels for enteric diseases were, therefore, structured as subregional panels, where some experts only provided estimates for a single subregion, whereas others provided estimates for all 14 subregions. The experts were free to decide the subregions for which they provided their judgments. Each panel elicited judgments for nine major enteric pathogens.

The decision to use global or multiregional panels implied potentially heavy time demands on expert study participants. For each hazard, experts were asked to provide a central estimate of the percentage of all cases of illness caused by the hazard that were due to a particular exposure route

[1] Use of the U.N. Food and Agricultural Organizations' Global Environmental Monitoring System Food Consumption Cluster Diet regions (GEMS regions) was considered and rejected. 2012 GEMS regions divide the world into 17 regions (Fig. 1b). GEMs regions are based on similarities in dietary patterns. Many factors in addition to diet affect exposure to foodborne hazards. Basic sanitation conditions, climate, endemic biota, animal husbandry and cropping practices, food processing, the prevalence of refrigeration, industrial pollution, food sourcing, the sophistication of supply chain management, the strength of regulatory regimes, and the effectiveness of regulatory enforcement all affect food contamination levels. Many GEMS 2012 regions cluster countries that are quite diverse in terms of factors that are likely to influence food safety and also cluster countries from different WHO regions. For example, Uruguay, Iceland, and Australia are clustered in a single GEMS region. Because of these factors, a decision was reached not to use GEMS regions.

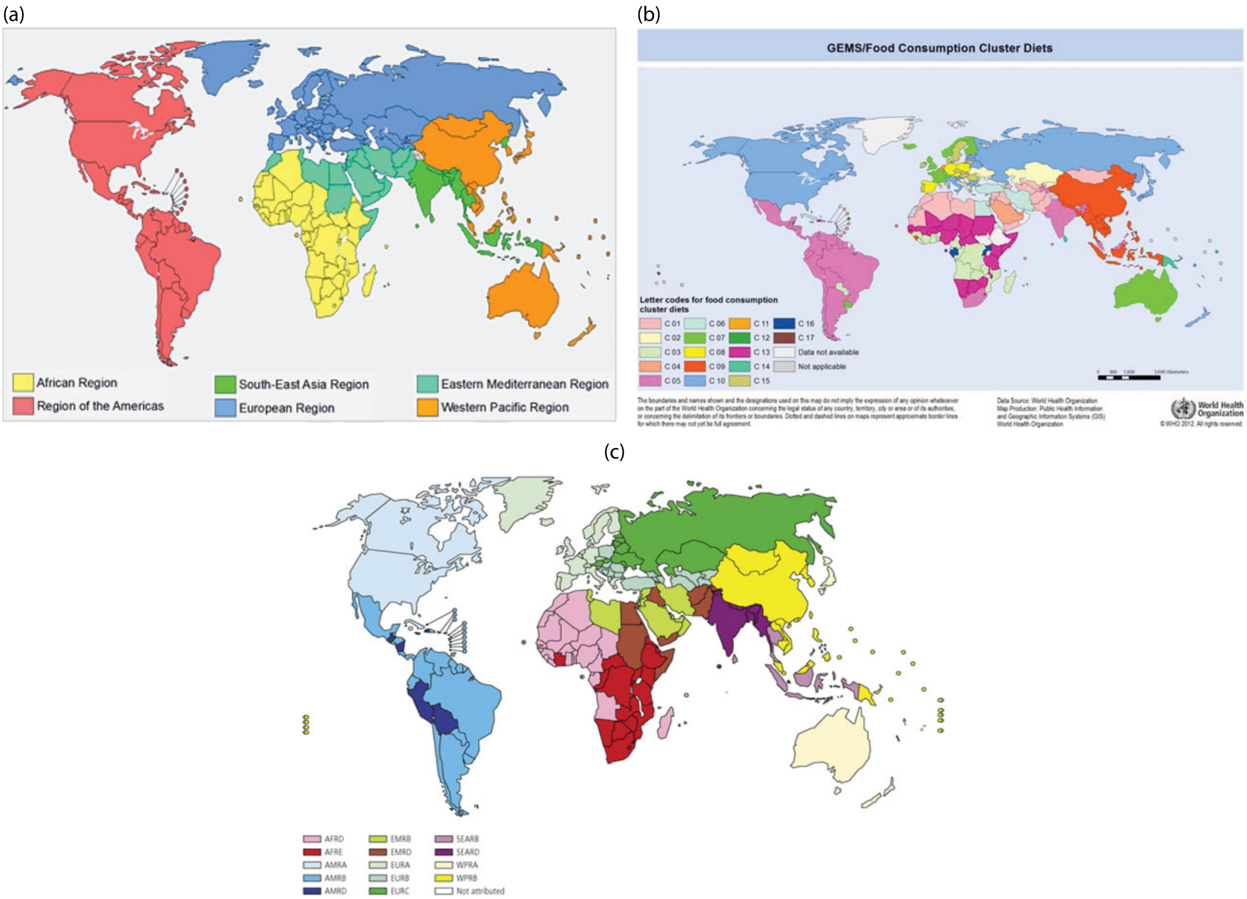

Fig. 1. (a) WHO regions.
Source: WHO, Definition of Region Groupings. Available at: http://www.who.int/healthinfo/global_burden_disease/definition_regions/en/, Accessed July 28, 2014.
(b) Global Environment Monitoring System (GEMS) regions.
Source: WHO, Global Environment Monitoring System—Food Contamination Monitoring and Assessment Programme (GEMS/Food): GEMS/Food Cluster Diets. Available at: http://www.who.int/foodsafety/chem/gems/en/index1.html, Accessed July 25, 2014.
(c)Global Burden of Disease subregions.
Source: WHO. Quantifying Environmental Health Impacts: Subregional Country Groupings for the Global Assessment of Disease Burden. Available at: http://www.who.int/quantifying_ehimpacts/global/ebdcountgroup/en/, Accessed July 25, 2014.

(Fig. 2). For each of these central estimates, experts were asked to provide a 90% credible interval (Fig. 2). They were asked to fill in a matrix like that shown in Fig. 2 for primary exposure pathways and another for foodborne exposure pathways for each pathogen. As shown in Table I, there are seven possible primary exposure routes and 14 possible food exposure routes for each of the 14 subregions. This meant that, for each pathogen, an expert could potentially be asked to provide 294 central estimates together with their corresponding credible intervals. Respondents tend to be expert in a class of hazards, for example, protozoan pathogens. As a result, given the limited number of experts on different hazards, the study needed to have some experts provide estimates for multiple pathogens (Table II); for example, experts on protozoa provided estimates for three different protozoan pathogens. This implied an unwieldy demand on experts' time. For example, some experts needed to provide estimates for nine enteric pathogens in 14 subregions (Table II).

Study researchers drew on the expertise of FERG members to narrow the scope of the elicitation by eliminating primary pathways or foods that scientific research and professional experience suggested were highly unlikely to be exposure routes. For example, this exercise resulted in excluding soil

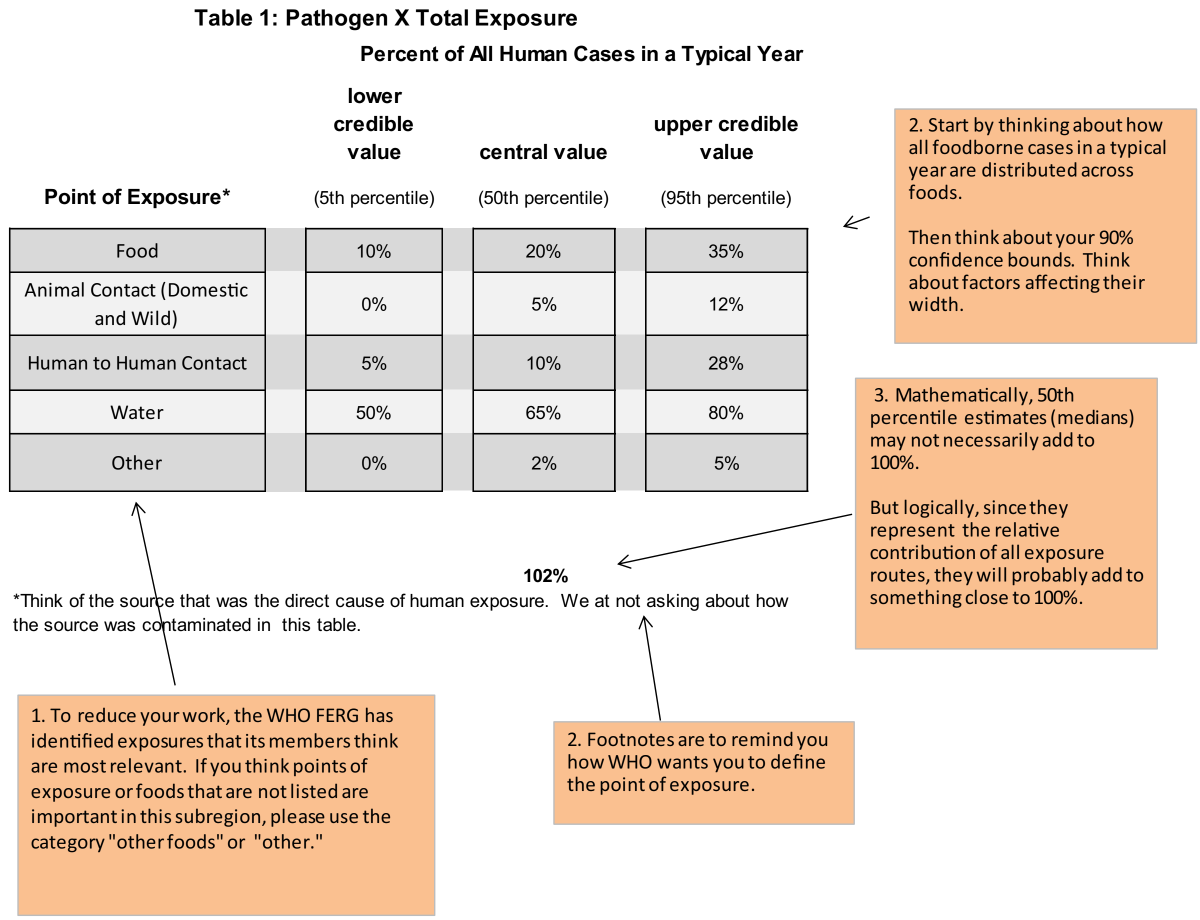

Fig. 2. Illustration of primary source attribution task for a hypothetical pathogen included in instructions for experts in the WHO Global Burden of Foodborne Diseases source attribution elicitation.

and air as possible exposure routes for most of the diarrheal enteric pathogens. Had more time and resources been available, it might have been possible in some cases to use systematic reviews to inform these decisions. But in most cases, there simply was not enough scientific literature for systematic reviews to be useful. This paring of exposure pathways based on FERG member recommendations substantially reduced the number of distributions that experts were asked to provide (Table III).

Experts were identified through iterative peer nomination, a method drawn from social network sampling research.[29,31] Care was taken to contact a diverse set of individuals and organizations as starting points in the nomination process to minimize bias. Starting points included nominations by leadership of scientific societies involved in the study of foodborne disease and food safety management, WHO regional offices working on food safety, and leading epidemiologists around the world working in fields related to foodborne disease. The nature and scope of the study was explained to the experts identified in this manner. These experts were then asked to nominate peers who they felt would have the requisite expertise to participate in the study. Final selection of expert panelists was made by the WHO in consultation with the leadership of the FERG using criteria for study requirements provided by the research team and standard WHO processes for screening experts' credentials. Evidence of expertise included peer recognition, research and publication, and documented professional experience in the field and food safety management. Attention was also paid to whether membership on each panel represented the needed geographic scope as well as a diversity of professional perspectives on foodborne disease and management. One hundred and three experts of the 299 identified through peer nomination

Table II. Structure of Panels for WHO Expert Elicitation on Exposures Leading to Foodborne Diseases

Panels	Number of WHO Subregions	Individual Hazards Included in Panel
Ascaris	14	*Ascaris lumbricoides*
Brucella	14	*Brucella* spp.
Echinococci	14	*E. granulosus*, *E. multilocularis*
HAV	14	Hepatitis A virus
Protozoa	14	*Cryptosporidium* spp., *Entamoeba histolytica*, *Giardia intestinalis*
Toxoplasma	14	*Toxoplasma gondii*
Other enterics (low mortality rate subregions)	3	*Campylobacter* spp., Enteropathogenic *E. coli* (EPEC), Enterotoxogenic *E. coli* (ETEC), Shiga-toxin producing *E. coli* (STEC), *Salmonella* nontyphoidal, *Salmonella* typhoidal, Norovirus, *Shigella*, *Vibrio cholera*
Other enterics (mid and high mortality rate subregions)	11	*Campylobacter* spp., Enteropathogenic *E. coli* (EPEC), Enterotoxogenic *E. coli* (ETEC), Shiga-toxin producing *E. coli* (STEC), *Salmonella* nontyphoidal, *Salmonella* typhoidal, Norovirus, *Shigella*, *Vibrio cholera*
Heavy metals	14	Lead

Note: Some experts participated in multiple panels.

were invited and agreed to participate in the elicitation. Of these, 78 were reached for in-person calibration question interviews, and 72 returned completed target-question questionnaires.

2.4.3. Development of Calibration Questions

Cooke's classical model uses responses to calibration questions to aggregate expert judgments. The purpose of the calibration questions is to assess the individual experts' ability to give statistically accurate probability assessments and to assess how informative their judgments tend to be relative to an uninformative background distribution. This is done by asking experts to provide median and 90% credible interval judgments for a set of 9–12 parameters/calibration questions. Experts should not know the answers to the calibration questions with certainty at the time they answer them. Ideally, calibration questions ask about future events for which data will be available after the calibration questions are answered, but before the answers are analyzed. This can be a difficult condition to meet and, therefore, calibration questions are often based on existing data or research, including systematic reviews, that is unavailable to experts at the time they complete the calibration questions.[32–34]

It is generally recommended that calibration questions be substantively related to the expert elicitation questions. In this case, the choice of calibration questions was guided by the need to represent the geographic variability in the way hazards are generated and the multidisciplinary nature of expertise about these processes. Experts with diverse backgrounds, including epidemiology, microbiology, parasitology, toxicology, veterinary and human medicine, and animal and food science participated in the study. The experts also had expertise in different geographic regions. Questions were drawn from a range of disciplines and regions in order not to disadvantage one expert relative to another. Where available, questions were based on global data sets; for example, questions were included based on WHO and FAO data on food consumption patterns, water and sanitation, and under-5 mortality.[35–37] Disease incidence data, available only for high-income countries, were also used to develop calibration questions. Individual experts tended not to be equally familiar with all of these data sources and familiarity varied by area of specialization. Questions based on systematic reviews were particularly useful in assuring that relevant disciplines and regions were covered by calibration questions. For example, it was possible to include questions relevant to sub-Saharan Africa, where published disease data are scarce, by drawing on systematic reviews and other scientific studies.

2.4.4. Administration of the Elicitation

Past elicitations using the Cooke classical model have generally been conducted in person by the studies' principal investigators (PIs). The global scope of this study and time and budgetary constraints made it impossible to use in-person elicitation or even elicitation only by experienced PIs. This is likely to be a constraint in other global expert elicitations. In this study, advanced Ph.D.

Table III. Possible and Actual Number of Exposure Routes for Each Biological Hazard in the WHO Expert Elicitation on Exposures Leading to Foodborne Diseases

		Number of Exposure Routes per Hazard (Maximum and Actual)		
		Primary exposure[a]	Food exposure	Other relationships
	Maximum possible exposure routes	7	14	4
Panel	Actual exposure outes elicited			
Ascaris	*Ascaris lumbricoides*	6	3	0
Brucella	*Brucella* spp.	5	5	0
Echinococci	*E. granulosus*	6	3	0
	E. multilocularis	6	3	0
HAV	Hepatitis A virus	4	0	0
Protozoa	*Cryptosporidium*	5	4	0
	Entamoeba histolytica	4	3	0
	Giardia intestinalis	5	3	0
Toxoplasma	*Toxoplasma gondii*	5	9	0
Other Enterics	*Campylobacter* spp.	6	8	4
	EPEC	5	0	0
	ETEC	5	0	0
	Norovirus	4	0	0
	Salmonella nontyphoidal	6	14	4
	Salmonella typhoidal	4	0	0
	Shigella	5	0	0
	STEC	6	7	4
	Vibrio cholera	4	0	0

[a]Primary exposure routes distinguish between foodborne and other major sources of exposure to hazards that can be foodborne.

students and postdoctoral fellows were trained to conduct the elicitations. Elicitations were conducted using one-on-one SKYPE or telephone calls. An effort was made to control experts' use of outside information in answering calibration questions by explaining the purpose of the questions and therefore the importance of not consulting references when answering the questions, and by having experts answer the calibration questions during a phone interview with a trained facilitator.

After the calibration questions were completed, each expert was given an elicitation instrument with which to provide judgments about exposure distributions of interest. Experts were given four weeks to develop their assessments of distributions of interest and were instructed that they could refer to any source they felt relevant while doing so. While experts were allowed to discuss the exercise with colleagues, they were reminded that the study needed their own best individual judgment. Facilitators maintained email communication with experts during the entire elicitation process. This system of administration worked well for completion of calibration questions and for responding to questions of clarification from experts. Greater difficulty was faced in encouraging experts to complete the elicitation, but 72 of the 78 experts who completed calibration interviews also returned completed target question instruments.

3. DISCUSSION

The organizers of the workshop on which this special issue is based posed a series of questions about lessons on research synthesis practice that could be drawn from workshop participants' experience and case studies. First, the organizers asked for criteria to evaluate when different research synthesis methods should be used. In the WHO global burden of foodborne disease initiative, systematic reviews were used in estimating the incidence and burden of specific diseases where there was sufficient depth in the published literature, but not for source attribution where there was little literature. Expert elicitation was used to provide source attribution only when other, more conventional options, including primary data analysis and risk assessment modeling, were infeasible. This was the case for most biological hazards. For some biological hazards it was possible to determine their primary and food exposure routes based on their biology/life cycle. Development of methods for food source attribution has been challenging at a national level and data requirements made it impossible in most cases to use conventional research or systematic reviews and meta-analysis to develop consistent regional estimates in a global study. Expert elicitation provided a needed alternative.

Second, workshop organizers asked for reflections on the strengths and limitations of research synthesis methods. In the WHO FERG setting, the flexibility of expert elicitation was a strength that allowed development of consistent estimates across hazards and WHO regions. Expert elicitation is particularly well suited to situations where professional experience and judgment can be used to reason about relationships where conventional empirical data are incomplete. In the case of foodborne disease source attribution, a wide range of factors converge to influence the likelihood that food exposure caused illnesses. These include the immunological health of the population, food consumption patterns, and a variety of food production and handling practices that could contribute to food contamination. Another strength of expert elicitation is that, by providing not only a central estimate, but also uncertainty bounds, it provides an indication of where uncertainty about parameter values is greatest. This is useful input into decisions about where greater investment in data collection or research is most needed.

A major limitation for the use of expert elicitation is its unfamiliarity to many scientists. Both conducting and using expert elicitation depends on the acceptance of scientists. Scientists are trained to analyze empirical data and may not appreciate the extent to which conventional science also relies on judgment. Where it is appropriate to use expert elicitation, the choice is not generally whether to rely on judgment or data, data should be used if they are available, of reasonable quality, and there is adequate time to analyze them. The choice is usually between the conventional practice of individual modeling teams making and justifying assumptions in model development versus more formal elicitation of judgments from other scientists with the best available expertise on parameters of interest. The use of formal, transparent, peer-reviewed methods that draw on best practices from related fields is critical to building the credibility of expert elicitation among scientists.

Finally, the organizers of the workshop also asked participants to draw on their experience to identify critical methodological research needs. The FERG source attribution expert elicitation used a formal expert elicitation method that has a significant track record. Our experience with the method and its application to a global management problem points to four areas that would benefit from more formal research. First, many expert elicitation methods elicit judgments from individual experts. There is a growing literature examining how best to aggregate individual expert judgments, but further research is needed comparing the performance of alternative aggregation methods and exploring ways of validating aggregation methods.[38–49] Second, specific to Cooke's classical model, more formal research on the design of calibration questions could be useful. Third, there is a need for research to develop more practical, reliable methods of eliciting expert judgment remotely. This need will only grow with the increasing need to manage risks globally. Baker *et al.* (2014) provide one of the few studies formally comparing in-person and self-administered elicitation methods.[50] Finally, it may be useful to give thought to what validation means in the context of expert elicitation. By its nature, expert elicitations tend to be conducted in areas where it is difficult to collect data and thus external validation may not be possible. Internal validity therefore becomes more important. Both tests for external validation, where possible, and development of criteria for internal validity would be useful.

4. CONCLUSION

The growing need for management of risks at a global scale has increased demand for risk information that is regionally or nationally specific, but globally comparable. Large global modeling efforts, like the WHO Global Burden of Foodborne Diseases Initiative, inevitably face significant challenges due to incomplete or inconsistent data. Geographical disparities in data collection and research can lead to very uneven availability and quality of data in different regions of the world, making it difficult to compare local, national, or regional conditions around the globe. The use of expert judgment is often unavoidable. Using formal expert elicitation rather than relying on informal judgments of individual modelers can help enhance the reliability and credibility of analysis of risks that are global in scope.

This study found that it was feasible to adapt formal expert elicitation methods to the problem of providing regional estimates worldwide. Expert elicitation provides peer-reviewed methods of eliciting and aggregating judgments on needed parameters from the best available scientific experts as an alternative to relying wholly on assumptions made by modeling teams. By providing distributions rather than point estimates for needed parameters, expert elicitation can provide estimates of how uncertainty about the parameters varies regionally around the globe. Use of formal, reproducible expert elicitation methods can bring a needed transparency to informing data gaps in global risk analysis modeling efforts.

ACKNOWLEDGMENTS

The views expressed in this article are those of the authors and do not necessarily represent the official views, decisions, or policies of the World Health Organization, U.S. Centers for Disease Control and Prevention, the U.S. Department of Agriculture, or other institutions listed.

This study was commissioned and paid for by the World Health Organization. Copyright in the original work on which this article is based belongs to WHO. The authors have been given permission to publish this article.

The authors are very grateful to the specialists who contributed their expertise and time to this study. We also thank our facilitators, who conducted the expert interviews and collected the data: Charlene Bayemi, Anca Hanea, Taher Jamshidi, Vikram Rao, Marshall Kuypers, Alessandra Lafranconi, Anuj Jindal, Sarvnaz Modaressi, Lohendy Muñoz, Jessica Kratchman, and Fjordor Melnikov. We would also like to acknowledge the assistance of the WHO Secretariat over the life of the FERG initiative, particularly Amy Cawthorne, Claudia Stein, and Yuki Minato.

REFERENCES

1. Robinson L, Hammitt J. Introduction to the special series on research synthesis: A cross-disciplinary approach. *Risk Analysis*, 2015; 35(6):963–970.
2. World Health Organization. WHO initiative to estimate the global burden of foodborne diseases: First formal meeting of the foodborne disease burden epidemiology reference group. Geneva (Switzerland) Nov 26–28, 2007, 2008.
3. Havelaar A, Cawthorne A, Angulo F, Bellinger D, Corrigan T, Cravioto A, Gibb H, Hald T, Ehiri J, Kirk M, Lake R, Praet N, Speybroeck N, de Silva N, Stein C, Torgerson P, Kuchenmüller T. WHO initiative to estimate the global burden of foodborne diseases. Lancet, 2013; 381:S59.
4. Kuchenmüller T, Hird S, Stein C, Cramarz P, Nada A, Havelaar A. Estimating the global burden of foodborne diseases—A collaborative effort. Eurosurveillance, 2009; 14(18):1–4.
5. Scallan, E, Hoekstra R, Angulo F, Tauxe R, Widdowson M, Roy S, Jones J, Griffin P. Foodborne illness acquired in the United States—Major pathogens. Emerging Infectious Disease, 2011; 17(1):7–15.
6. Scallan E, Griffin P, Angulo F, Tauxe R, Hoekstra R. Foodborne illness acquired in the United States—Unspecified agents. Emerging Infectious Disease, 2011; 17(1):16–22.
7. Hoffmann S, Batz M, Morris JG Jr. Annual cost of illness and quality-adjusted life year losses in the United States due to 14 foodborne pathogens. Journal of Food Protection, 2012; 75(7):1291–1302.
8. Stein C, Kuchenmüller T, Hendrickx S, Prüss-Ustün A, Wolfson W, Engels D, Schlundt J. The global burden of disease assessments—WHO is responsible? PLoS Neglected Tropical Diseases, 2007; 1(3):e161.
9. World Health Organization. WHO initiative to estimate the global burden of foodborne diseases: A summary document, 2008.
10. U.N. Food and Agricultural Organization and World Health Organization (FAO/WHO). The role of food safety in health and development. Geneva, Switzerland: WHO, 1984. WHO Technical Report Series 705. Available at: http://apps.who.int/iris/bitstream/10665/38709/1/WHO_TRS_705.pdf?ua=1, Accessed July 25, 2014.
11. Hoffmann S, Harder W. Food safety and risk governance in globalized markets. Health Matrix: Journal of Law-Medicine, 2010; 20:5–54.
12. Devleesschauwer B, Havelaar A, Maertens de Noordhout C, Haagsma J, Praet N, Dorny P, Duchateau L, Torgerson P, Van Oyen H, Speybroeck N. Calculating disability-adjusted life years to quantify burden of disease. International Journal of Public Health, 2014; 59:565–569.
13. Murray C. Quantifying the burden of disease: The technical basis for disability-adjusted life years. Bulletin World Health Organization, 1994; 72(3):429–445.
14. Salomon J, Vos T, Hogan D, et al. Common values in assessing health outcomes from disease and injury: Disability weights measurement study for the global burden of disease study 2010. Lancet, 2012; 380(9859):2129–2143.
15. Carabin H, Ndimubanzi P, Budke C, Nguyen H, Qian Y, et al. Clinical manifestations associated with neurocysticercosis: A systematic review. PLoS Neglected Tropical Diseases, 2011; 5(5):e1152.
16. Fischer Walker C, Sack D, Black R. Etiology of diarrhea in older children, adolescents and adults: A systematic review. PLoS Neglected Tropical Diseases, 2010; 4(8):e768.
17. Fürst, T, Keiser J, Utzinger J. Global burden of human foodborne trematodiasis: A systematic review and meta-analysis. Lancet Infectious Disease, 2012; 12(3):210–221.
18. Ndimubanzi P, Carabin H, Budke C, Nguyen H, Qian Y-J, et al. A systematic review of the frequency of neurocyticercosis with a focus on people with epilepsy. PLoS Neglected Tropical Diseases, 2010; 4(11):e870.
19. World Health Organization. FERG strategic planning meeting and kick-off event of the pilot country studies in Durres, Albania, Nov 7–10, 2011, 2011. Available at: http://www.who.int/foodsafety/foodborne_disease/ferg_albania/en/, Accessed July 25, 2014.
20. European Food Safety Agency. Overview of the procedures currently used at EFSA for the assessment of dietary exposure to different chemical substances. EFSA Journal, 2011; 9(12):2490–2523.
21. Pires S. Assessing the applicability of currently available methods for attributing foodborne disease to sources, including food and food commodities. Foodborne Pathogens and Disease, 2013; 10(3):206–213.
22. U.S. Centers for Disease Control and Prevention (CDC). Interagency Food Safety Analytics Collaboration (IFSAC), 2014. Available at: http://www.cdc.gov/foodsafety/ifsac/index.html, Accessed July 25, 2014.
23. Aspinall W. A route to more tractable expert advice. Nature, 2010; 463:294–295.
24. Van der Fels-Klerx H, Cooke R, Nauta M, Goossens L, Havelaar A. A structured expert judgment study for a model of *Campylobacter* transmission during broiler-chicken processing. Risk Analysis, 2005; 25(1):109–124.

25. Havelaar A, Vargas Galindo A, Kurowicka D, Cooke R. Attribution of foodborne pathogens using structured expert elicitation. Foodborne Pathogens and Disease, 2008; 5(5):649–659.
26. Flandoli F, Giorgi E, Aspinall W, Neri A. Comparison of a new expert elicitation model with the classical model, equal weights and single experts, using a cross-validation technique. Reliability Engineering & System Safety, 2011; 96(10):1291–1310.
27. Cooke R. Experts in Uncertainty: Opinion and Subjective Probability in Science. New York: Oxford University Press, 1991.
28. Cooke R, Goossen L. TU Delft expert judgment data base. Reliability Engineering and System Safety, 2007; 93:657–674.
29. Goodman L. Snowball sampling. Annals of Mathematical Statistics, 1961; 31(1):148–170.
30. Ezzati M, Lopez A, Rodgers A, Vander Hoorn S, Murray C, and the Comparative Risk Assessment Collaborating Group. Selected major risk factors and global and regional burden of disease. *Lancet*, 2002; 360:1347–1360.
31. Atkinson R, Flint J. Encyclopedia of Social Science Research Methods. Thousand Oaks, CA: SAGE Publications, Inc., 2004.
32. Bamber J, Aspinall W. An expert judgement assessment of future sea level rise from the ice sheets. Nature Climate Change, 2013; 3:424–427.
33. Baxte P, Aspinall W, Neri A, Zuccaro G, Spence R, Cioni R, Woo G. Emergency planning and mitigation at Vesuvius: A new evidence-based approach. Volcanology and Geothermal Research, 2008; 178(3):454–473.
34. Tyshenko M, ElSaadany S, Oraby T, Darshan S, Catford A, Aspinall W, Cooke R, Krewski D. Expert judgement and re-elicitation for prion disease risk uncertainties. International Journal of Risk Assessment and Management, 2012; 16(1):1466–8297.
35. U.N. Food and Agricultural Organization (FAO). Food balance sheets. Available at http://faostat.fao.org/site/354/default.aspx, Accessed July 25, 2014.
36. World Health Observatory Data Repository. Under-five mortality due to diarrhea. Available at http://apps.who.int/gho/data/node.main, Accessed 24 June 14, 2014.
37. WHO/UNICEF Joint Monitoring Programme for Water Supply and Sanitation. Progress on drinking water and sanitation: Joint Monitoring Programme 2012 update. New York: WHO/UNICEF, March 2012. Available at http://www.who.int/water_sanitation_health/publications/2012/jmp_report/en/, Accessed October 29, 2014.
38. Eggstaff J, Mazzuchi T, Sarkani S. 2014. The effect of the number of seed variables on the performance of Cooketicle/pii/S0951832013002251" ort/en/" d at http://apps.who.int/gh 2014; 121:72–82.
39. Christopher W, Karvetski C, Olson K, Mandel D, Twardy C. Probabilistic coherence weighting for optimizing expert forecasts. Decision Analysis, 2013; 10(4):305–326.
40. Jose V, Grushka-Cockayne Y, Lichtendahl K Jr. Trimmed opinion pools and the crowd's calibration problem. Management Science, 2013; 60(2):463–475.
41. Albert I, Donnet S, Guihenneuc-Jouyaux C, Low-Choy S, Mengersen K, Rousseau J. Combining expert opinions in prior elicitation. Bayesian Analysis, 2012; 7(3):503–532.
42. Clemen R, Winkler R. Aggregating probability distributions. Pp. 154–176 in Edwards W, Miles R Jr., von Winterfeldt D, (eds). Advances in Decision Analysis: From Foundations to Applications. New York: Cambridge University Press, 2007.
43. Lin S, Bier V. A study of expert overconfidence. Reliability Engineering and System Safety, 2008; 93(5):711–721.
44. Lin S, Huang S. Effects of overconfidence and dependence on aggregated probability judgments. Journal of Modelling in Management, 2012; 7(1):6–22.
45. Lin S-W, Cheng C-H. Can Cooke's Model Sift Out Better Experts and Produce Well-Calibrated Aggregated Probabilities? Department of Business Administration, Yuan Ze University, Chung-Li, Taiwan Proceedings of the 2008 IEEE IEEM, 2008.
46. Cooke R. Pitfalls of ROAT cross validation comment on effects of overconfidence and dependence on aggregated probability judgments. Journal of Modelling in Management 2012; 7(1):20–22.
47. Clemen R. Comment on Cooke's classical method. Reliability Engineering & System Safety, 2008; 93(5):760–765.
48. Cooke R. Response to comments, special issue on expert judgment. Reliability Engineering & System Safety, 2008a; 93(5):775–777.
49. Cooke R, lson A. Cross Validation for Structured Expert Judgment with the Classical Model with Supplementary Online Material, 2015. Discussion Paper Available at: http://rogermcooke.net.
50. Baker E, Bosetti V, Jenni K, Ricci E. Facing the Experts: Survey Mode And Expert Elicitation, Milan, Italy; Fondazione Eni Enrico Mattei Note di Lavoro, Jan 2014. Available at: http://www.feem.it/getpage.aspx?id=6046&sez=Publications&padre=73 October 27, 2014.