Bounding Analysis of Lung Cancer Risks Using Imprecise Probabilities*

MINH HA-DUONG
Centre National de la Recherche Scientifique, France

ELIZABETH CASMAN
Carnegie Mellon University, USA

M. GRANGER MORGAN
Carnegie Mellon University, USA

Abstract

For cancers with more than one risk factor, the sum of probabilistic estimates of the number of cancers attributable to each individual factor may exceed the total number of cases observed when uncertainties about exposure and dose-response for some factors is high. In this study we outline a method to bound the fraction of lung cancer fatalities not attributed to specific well-studied causes in which available data and expert judgment are used to attribute portions of the observed lung cancer mortality to known causes such as smoking, residential radon, and asbestos fibers. An upper bound on the residual risk due to other causes is then inferred using a coherence constraint on the total number of deaths, a maximum uncertainty principle, and imprecise probabilities.

Keywords

bounding analysis, lung cancer, belief functions, assessment methods, medicine

1 Introduction

Usually, the health risk of exposure to an environmental contaminant is calculated using a "front-to-back" procedure, which involves estimating toxic releases, modeling environmental and physiological transformations, and then employing exposure models and dose-response functions, see for example [6]. That methodology works best when the relevant science is well developed; however, when

*This work has been supported by CNRS of France, and by the National Science Foundation under grants SES-0216897 and SRB-9521914. A similar paper has been submitted for publication in the journal *Risk Analysis*.

Well characterized factors	Less well characterized factors
Cigarette smoking	Occupational exposures:
Passive smoking	Asbestos
Indoor radon	Arsenic
	Chromates
	Chloromethyl ethers
	Diesel exhaust
	Nickel
	Polycyclic aromatic hydrocarbons (PAHs)
	Ambient air pollution

Table 1: Examples of environmental risk factors for lung cancer

there are several *risk factors* (as the expression is used in the epidemiology literature), and uncertainty about some of the science is large, such a procedure can lead to estimates for the numbers of cancers attributable to the various factors that, summed, exceed the total number of cases actually observed.

Morgan [12] argued that methods of bounding analysis could be used for environmental risk analysis. For health risks with multiple external causes, the available knowledge constrains the magnitude of the poorly characterized risks. If most risks were known with precision, this would be a simple subtraction problem. However disease risks from environmental causes are often estimated from models or inferred from studies involving limited numbers of subjects and inconsistent notions of controls or have other methodological problems that contribute to the uncertainty of the results. It is common to see the central tendencies of such risk estimates expressed as ranges, especially when there are competing plausible models. Sometimes the sum of the individual risks exceeds the total risk. How to quantify and bound the residual "unclaimed" risk is the subject of this paper.

Using lung cancer mortality from environmental factors as an illustrative example, this paper presents a method for bounding the remaining uncertainty when only some of the risk factors are well characterized. The result is an upper bound on the mortality that can be attributed to all other, less well-characterized factors. Some of the major environmental risk factors for lung cancer are shown in Table 1. "Well characterized" here means that population-wide longitudinal attributional studies exist.

In the method presented, expert judgment is used to attribute a portion of the observed cancers to known causes such as smoking, radon and asbestos. Information about the risks from unspecified causes is inferred using a coherence constraint on the total number of deaths, and a principle we term maximum uncertainty.

Our method builds upon the work of Walley [23, chapter 4]. Mathematically, this is an application of Smets' Transferable Belief Model [20], which was de-

veloped to solve some paradoxes in combining expert opinion in the theory of evidence [19]. We elicit information about a finite set of variables (risk factors for cancer) and represent this information as constraints on a linear programming problem involving a convex family of probabilities. We invoke the maximum unspecificity criterion in order to estimate the upper bound for the less well-studied members of the set.

Ours is not the first combination of linear programming, expert elicitation, and imprecise probabilities. Lins combined these elements [10] to assess prior probabilities for a single continuous parameter.

The paper is organized as follows. Section 2 presents the conceptual model, which is an application of the mathematical Transferable Beliefs Model to risk assessment. Based on this, Section 3 discusses our method to elicit and validate expert opinion using a maximum unspecificity criterion. From our reading of the literature, we then provide a tentative attribution among the causes (because the expert elicitation phase of this project is currently incomplete), and in Section 4 illustrate the method with a numerical application.

2 Model

2.1 Multiple pollutants may cause lung cancer

Let N denote the magnitude of the health end-point, in this case, the total annual number of lung cancer deaths. Let Ω denote the set of all possible causes of lung cancer deaths. For example, $\Omega = \{C, \quad R, A, X\}$ where C means tobacco smoke primarily from cigarettes, R means indoor exposure to radon, A means asbestos and X is the group of all other more poorly understood environmental factors of interest.

The model assumes that N is readily observable and therefore known with precision. While this is not strictly true in the case of lung cancer [3, 2] the assumption is not limiting, since the results of the method can be stated in percentage terms and then applied to a range of possible numerical values of N. We also assume exposure to be binary, which is of course not true, but the assumption is consistent with the exposure definitions used in the supporting epidemiological studies. With these two assumptions, each death can be linked to zero or more possible causes in Ω. Most lung cancer deaths are caused by smoking alone, but there are synergistic cases in which more than one cause is involved, such as smoking and radon.

Figure 1 shows one way to subdivide N by causes that includes synergistic effects. We denote the number of deaths linked to cause s as $n(s)$, where s is any subset of Ω. In our example we consider four possible causes in Ω, so there could be sixteen ($= 2^4$) possible s, but to simplify the analysis and to be consistent with the cancer literature, we will consider only the two-factor interactions involving

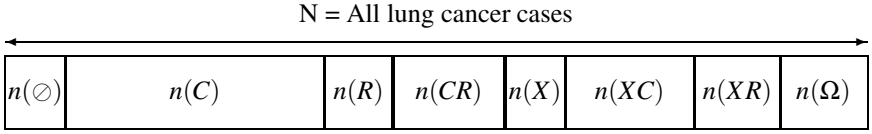

Figure 1: The basic statistic n, simplified to include only the risk factors cigarettes (C), radon (R), and all other causes (X). N is the total number of lung cancer fatalities. n is the number of fatalities attributable to each risk factor, or combination of factors. $n(\oslash)$ is the background number of lung cancer deaths that would occur absent all the various risk factors. $n(\Omega) = n(CRX)$, those cases for which no risk factor can be excluded.

cigarette smoke.

To adopt a more precise and cautious definition, $n(s)$ is the number of cases not caused by pollutants not in s. This implies that causes not in s are known to be non-contributing to that lung cancer. For deaths in $n(s)$, any cause in s may have caused the lung cancer, but which one is uncertain and there may have been synergies.

Our intuitive interpretation for this definition of "ambiguous causality" is that $n(s)$ represents the number of cases that were exposed to the possibly multiple risk factors in s.

The number of lung cancer deaths where all causes of Ω have been positively excluded is $n(\oslash)$ shown to the left of the bar in Figure 1. Cases that could not be linked to any pollutant in Ω are considered spontaneous lung cancer. It is important to underline that $n(\oslash)$ does not have the same status as $n(X)$, which will be deduced as a residual. It corresponds to the background rate of lung cancer that occurs in a population without exposure to any environmental, dietary, occupational or other carcinogen.

The function n does not come from real data. Direct measurement of the basic statistic n is impossible, since exposure to a pollutant does not necessarily result in a cancer fatality and because retrospectively, lifetime exposures to the various carcinogens can only be roughly estimated. It is only a mathematical tool used in to support expert elicitation of consistent bounds, as discussed next.

2.2 Bounding the risk attributable to single and joint pollutants

The basic statistic n can be used to bound the number of cases attributable to smoking C as follows, where $\overline{n}(C)$ and $\underline{n}(C)$ denote the upper and lower bounds on $n(C)$, respectively:

- The lower bound is the number of cases attributed only to smoking (we

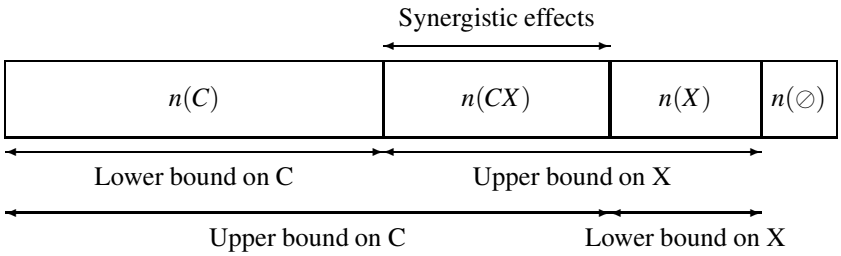

Figure 2: Upper and lower bounds on the number of lung cancer deaths attributable to C and X

lump both passive and active smoking together). That is $\underline{n}(C)$=$n(C)$.

- The upper bound is the number of cases exposed to smoke and possibly other factors. That is $\overline{n}(C) = n(C) + n(XC) + n(CR) + n(XCR) + n(CA) + n(XCA) + n(CRA) + n(XCRA)$ or:

$$\overline{n}(C) = \sum_{E} n(E) \text{ for all subsets } E \text{ of } \Omega \text{ containing } C$$

Figure 2 illustrates this definition of the upper and lower bounds of the number of lung cancer deaths attributable to X and C. For clarity the figure is drawn showing only two causes, with $\Omega = \{C, X\}$.

In epidemiologists' terms, the *attributable fraction* of pollutant C is the proportion of all cases that could be avoided if this pollutant were eliminated, denoted $af(C)$. The model suggests the following bounds for smoking attributable fraction:

$$\frac{\underline{n}(C)}{N}(1 - r_0) \leq af(C) \leq \frac{\overline{n}(C)}{N}(1 - r_0) \tag{1}$$

The lower bound accounts for the $1 - r_0$ share of spontaneous lung cancer cases in those cases exposed to cigarettes. The upper bound attributes all cigarette-exposed deaths to this factor.

For this paper we will assume that the background rate of lung cancer mortality in the U.S., r_0, is 3 deaths per 100 000 people. This background rate is the number of lung cancer deaths in the unexposed population divided by the unexposed population. Denoting p_C, p_R and p_A as the exposure probabilities of C, R, A; and T as the total population; assuming independence (meaning that people who smoke are no more or less likely to be exposed to radon or to asbestos):

$$r_0 = \frac{n(\oslash)}{(1 - p_C)(1 - p_R)(1 - p_A)T} \tag{2}$$

Consider now the bounds on deaths attributed to multiple synergistic causes. Denote these causes s, a subset of Ω, for example $s = CR$. For the lower bound on the number of deaths attributable to these causes acting jointly, we continue to adopt the number of cases exposed only to these causes, that is:

$$\underline{n}(s) = n(s) \tag{3}$$

And as the upper bound, we continue to adopt the number of cases exposed to s and possibly other factors, that is:

$$\overline{n}(s) = \sum_{E} n(E) \text{ for all subsets } E \text{ of } \Omega \text{ containing } s \tag{4}$$

This $\overline{n}$ corresponds to the commonality function in the Transferable Belief Model [20]. Bounds on the attributable fraction can be computed as in equation 1.

2.3 Unspecificity, a measure of uncertainty

Structurally, the only uncertainty in this bounding analysis model comes from the synergistic causes, because it is not possible to attribute the cancer to any one of these causes. Consider these two (of the three) extreme cases:

- If each death were attributed to exactly one cause, then there would be no uncertainty, and all lower bounds would coincide with their upper counterpart. We would have $n(C)+n(R)+n(A)+n(X) = N - n(\oslash)$. Note that since n is a positive function that sums up to N, this implies that $n(s) = 0$ for all other subsets.
- If no information were available, each death would be attributed to the synergy of all factors. We would have all the lower bounds at 0 and all upper bounds at N. Mathematically, this is $n(\Omega) = N$. Note that this constitutes a proper uninformative distribution: it is not the Bayesian uniform prior probability distribution on Ω. It represents the family of all probability distributions that can be defined on Ω.

Unspecificity is an numeric indicator that equals one in the first case, and in the second case equals the number of elements of Ω. It is the expected value of the number of elements of s with respect to the probability distribution $m(s) = \frac{n(s)}{N}$, that is:

$$U = \frac{n(C)+n(R)+n(A)+n(X)+2(n(CR)+n(RA)+\ldots)+3\ldots+4n(\Omega)}{N} \tag{5}$$

In this paper unspecificity is a kind of generalized cardinality, that specifies the number of alternatives. The reason for using this word is that when a death is

attributed to the synergy of k factors, it can be said that the unspecificity of this information is k. See [16] for an extensive discussion of this concept.

A lower unspecificity measure corresponds to better information, so a third extreme case needs discussion: unspecificity is zero when and only when $n(\oslash) = N$. This is the case when for all deaths, all non-spontaneous causes of Ω have been positively excluded. It means that all the substances in Ω are actually safe (with respect to lung cancer). This is the highest level of information achievable, to the point that it makes Ω irrelevant.

Regarding unspecificity as a measure of information allows to implement numerically the general principle of maximum uncertainty, also known as Laplace's principle of "raison insuffisante". The principle states that one should select the statistic that is the most unspecific, compatible with existing information. This is the principle that we use in the next section to estimate the bounds on the unknown cause, given information about all others.

3 Expert elicitation

3.1 Procedure

When we apply this procedure, we will elicit a set of judgments regarding $n(s)$ from a number of leading health scientists using methods previously developed for expert elicitation in domains in which there is considerable scientific evidence [13, 14, 15]

The results from an elicitation will be interpreted as linear constraints on n. These constraints determine a set $\mathcal{B}$ of basic statistics, that is a set of n that are all compatible with the expert's judgments. The most unspecific n in $\mathcal{B}$ is chosen to represent the expert judgment, according to the maximum uncertainty principle. This amounts to solving a linear program in a space with $2^{|\Omega|}$ dimensions.

Other ways of translating judgments into constraints are possible, for example using relative risk, but are not used in this introductory paper. Note that both quantitative and comparative judgments are possible, which may ultimately be important because some of the pollutants have been well studied, but we are interested in the less well-known pollutants.

In addition to elicited information, we impose these constraints:

- It is understood that all $n(s)$ are non-negative, summing up to N.
- Three-way interactions and higher are not allowed. That is, $n(s) = 0$ if s has 3 or more elements.

The constraint on three-way interaction is a zero-order approximation. We assume that that the number of deaths caused by multiple interactions are a very small number that can be neglected. In a more sophisticated approach, this assumption could be replaced by explicit considerations about causes interactivity

and independence. But there is little scientific empirical knowledge about these interactions.

3.2 Ensuring consistency

Maximizing unspecificity is possible only if $\mathcal{B}$ is not empty. This means that the different items of information given by the expert should be coherent with each other. For example, one could not allow the expert to say that the lower bound for C is 90 percent, and the lower bound for R is 20 percent at the same time, because that would exceed 100 percent. Walley has shown [23] that the coherence condition is:

$$\overline{af}(s_i) + \sum_{j \neq i} \underline{af}(s_j) \leq 1 \leq \underline{af}(s_i) + \sum_{j \neq i} \overline{af}(s_j)$$

The double inequality should hold for all causes i in $\{1, \ldots, |\Omega|\}$.

Besides mathematical consistency, it is also important to provide safeguards so that the expert can check that formal implications of the elicited n are consistent with its informal understanding of the problem. We propose two checks.

The first check on n is to make sure that the results in terms of bounds on relative risks and on interactions between pollutants make sense. The definition of relative risk for smoking cigarettes $rr(C)$, for example, is the lung cancer rate associated with exposure to tobacco smoke divided by the background lung cancer rate. Given exposure probabilities in the general population, we will assess the bounds on the relative risk for the various pollutants using the formula in [6, appendix C p. 229].

The second check on n is to make sure that the risk-ranking it implies makes sense. We will ask experts to rank risks during the elicitation process. The consistency of results will be assessed by comparing the partial order derived from n with the expert's *a priori* risk ranking.

Informally, this partial order says that the lung cancer risk related to R is not larger than the risk related to C when we know with certainty that R causes fewer lung cancer deaths than C. For example, one sufficient condition for this is that the lower bound on C is greater than the upper bound on R. But the mathematical definition of the natural partial order relation associated with a basic statistic n requires more explanations.

Let P denote a function such that $P(C)+P(R)+P(A)+P(X)=N$. It is a basic statistic with unspecificity one, describing an hypothetical world where each lung cancer death is attributed to one and only one cause. For such a P, the number of deaths caused by any set of causes s is $\sum_x P(x)$, for all causes x in s. We say that P is compatible with the basic statistic n if and only if for all s, that number respects the bound determined by n in the following way:

$$\forall s \subset \Omega, \sum_{x \in s} P(x) \leq \sum_{y, s \cap y \neq \oslash} n(y) \tag{6}$$

The right hand side of Equation 6 can be interpreted in the present model as the upper bound on the number of deaths related to the causes in s acting either jointly *or separately*. This function of s corresponds to the belief function in the Transferable Belief Model.

The heart of the problem is that P is hypothetical. Because there are interactions, more than one P is compatible with n. Denote $\mathcal{P}$ the family of all P compatible with n. The natural partial order is mathematically defined by:

$$R \preceq C \Leftrightarrow \forall P \in \mathcal{P},\ P(R) \leq P(C) \tag{7}$$

Numerically, this is determined by checking the sign of the minimum of $P(C) - P(R)$ under constraint 6. It is tractable to work with the full partial order, since there is at most $|\Omega|(|\Omega| - 1)/2$ comparisons. Assuming $|\Omega| = 7$ for example, there are no more than 21 information items, which can be presented naturally in the diagonal half of a table. Moreover, practically there will be fewer than 21 items, since not all risks can be compared. It is to be expected, for example, that some experts may prefer to find that some of the less-known risks are not comparable, because of missing scientific information.

4 Application

Our numerical simulations were performed using a *Mathematica* notebook[1]. The code directly implements matrix calculus for belief functions as outlined in [21]. This is the most straightforward method given that Ω remains small, but it would not scale well to tens of pollutants, since it involves square matrices with $2^{2|\Omega|}$ elements. For example, 10 pollutants implies storing in memory arrays with 1M numbers.

In our illustration Ω, the set of possible causes of lung cancer, consists of:

- C Smoking
- R Radon
- A Asbestos, glass wool, ceramic fibers
- X All other environmental risk factors

Based on our own review of the literature [6, 7, 4, 18, 22, and others] we have constructed a set of judgments attributing lung cancer deaths among the major causes, as the expert elicitations have not at this time been performed. We offer the following breakdown: Cigarette smoking combined with passive smoking accounts for 70 to 95 percent of lung cancer mortality; indoor radon exposures for 02 to 21 percent; asbestos, 1 to 5 percent.

[1] Available on the web at `http://www.andrew.cmu.edu/user/mduong`, or upon request, under the GNU General Public License.

Bounds	C	R	A	X
$\overline{af}$	95%	21%	5%	3.2%
$\underline{af}$	70%	02%	1%	0%
Exposure probability	45%	50%	5%	5%
$\overline{rr}$	43.2	1.53	2.05	1.66
$\underline{rr}$	6.19	1.04	1.20	1.

Table 2: Results of optimization: Upper and lower bounds on attributable fractions and relative risks

We used a 3% background rate [1, 5, 11, 9, 17]. With our assumptions on exposure probabilities, equation 2 implies that $n(\oslash) = 0.013N$.

The next table shows the implications for bounds of af and rr of the most unspecific imprecise probability distribution compatible with these constraints. The exposure probabilities needed to compute rr are exogenous: radon exposure is defined as living in a home with radon concentration at or above 25 Bqm^{-1}, and exposure to X is our estimate. The effect of this calculation on the bounds of rr would serve as a calibration/validation reference for the expert who may be more familiar with small sample studies than population effects, and might adjust his or her initial responses in light of seeing their mathematical implications.

This result attributes between 0 and 3.2 percent of lung cancer deaths to X, the group of unknown environmental pollutants. For the group of known and suspected lung carcinogens other than C, A and R, the risk analyst concludes that, *if one is confident in the bounds assigned to the well understood risk factors*, the sum of the effects of the other factors accounts for no more than 3.2% of total lung cancer mortality.

The implication for judging future risk assessments of members of X is that, if the assessment projects the lung cancer risk in the U. S. population from these pollutants to be in excess of 3.2% of the annual lung cancer mortality, then the assumptions of the model should be re-examined and the upper bound on the resulting estimate constrained.

5 Concluding remarks

5.1 Discussion

With less than ten pollutants, computing time is not a problem. Expert elicitation could be done interactively, solving for n after each expert's reply. This would allow the interviewer to point out and resolve inconsistency when there is no solution. But assuming that experts were willing to form judgments on a wider range

of pollutants, the curse of dimensionality can be addressed along the following lines. Rather than using matrix calculus, it is possible to use faster algorithms (namely the Fast Möbius transform) for belief function computations. If this is not enough, further simplifications can be made if additional assumptions on n, for example disallowing 3-way or higher interactions, are accepted.

The proposed method takes all information items provided by the expert with equal force. A potential advance of this research could be to ask experts to rank the reliability of each information item, or even to give an estimate of confidence for them.

Further research could deal with inter-expert validation, a question linked with the unresolved issue of judgment fusion. The Transferable Belief Model underlying this work offers a measure of contradiction between different sources of information: it reinterprets $n(\oslash)$, the number of spontaneous lung cancer deaths found when one combines the opinion of all experts. The problem is how to combine the experts.

Each expert's judgment determines a set $\mathcal{B}$ of coherent basic statistics. If the intersection of all these sets is non-empty, then experts agree on this intersection. The principle of maximum unspecificity can be used to form a group judgment.

If the intersection is empty, the experts contradict each other. Studying which information items cause the contradiction (which constraints make the LP infeasible) can identify the substantive sources of disagreement, and in that way inform both future research priorities as well as the decision-making process. How (or if) to fuse the judgments and quantify the degree of contradiction is still an active research question, see [8] for example.

5.2 Conclusion

This paper has proposed an application of the Transferable Belief Model [20] to estimate an upper bound on the number of lung cancers caused annually by the group of causes for which comprehensive longitudinal studies are lacking. Such a result is interesting from a risk management perspective, as it gives an indication of the level of effort control of these pollutants deserve.

This was done by attributing a portion of the observed cancers to known causes such as smoking, radon and asbestos, and then deducing information about the residual using maximum unspecificity. The critical aspects of this procedure are:

1. Uncertainty in the known causes is explicitly stated, using statements on upper and lower bounds.

2. Synergistic effects in the known causes are part of the framework.

3. Consistency between known causes and poorly understood agents is required. (As Figure 2 illustrates, it is the lower bound on smoking that mostly

constrains the upper bound on the residual.)

This paper presents the methodology. The results revealed by future expert elicitation will be the subject of another paper.

References

[1] ALAVANJA, M. C. R., BROWNSON, R. C., BENICHOU, J., SWANSON, C., AND BOICE, J. D. J. Attributable risk of lung cancer in lifetime nonsmokers and long-term ex-smokers (missouri, united states). *Cancer Causes and Control 6* (1995), 209–216.

[2] AMERICAN CANCER INSTITUTE. Cancer facts & figures. Tech. rep., American Cancer Society, Inc. Surveillance Research, 2003.

[3] ARCHER, V. E., AND LYON, J. L. Errors and biases in the diagnosis of cancer of the lung and their influence on risk estimates. *Medical hypotheses 54(3)* (2000), 400–407.

[4] AXELSON, O. Alternative for estimating the burden of lung cancer from occupational exposures – some calculations based on data from swedish men. *Scandinavian Journal of Work, Environment, and Health 28*, 1 (2002), 58–63.

[5] AXELSON, O., DAVIS, D. L., FORESTIERE, F., SCHNEIDERMAN, M., AND WAGENER, D. Lung cancer not attributable to smoking. *Annals of the New York Academy of Sciences 609*, 78 (1990), 165–178.

[6] BEIR VI, C. *Health effects of exposure to radon*. National Academy Press, 1999.

[7] CENTERS FOR DISEASE CONTROL AND PREVENTION. Smoking-attributable mortality, morbidity, and economic costs (SAMMEC): adult SAMMEC and maternal and child health (MCH) SAMMEC software, 2002.

[8] ISIF (INTERNATIONAL SOCIETY ON INFORMATION FUSION). *Fusion 2000 Conference* (Paris, 10–13 July 2000).

[9] KOO, L. C., AND HO, J. H. Worldwide epidemiological patterns of lung cancer in nonsmokers. *International Journal of Epidemiology 19*, Suppl. 1 (1990), s14–s23.

[10] LINS, G. C. N., AND DE SOUZA, F. M. C. A protocol for the elicitation of prior distributions. In *Second International Symposium on Imprecise Probability and their Applications* (New York, USA, 26–29 June 2001), Cornell University.

[11] LYON, J. L., GARDNER, J. W., AND WEST, D. W. Cancer in utah: risk by religion and place of residence. *J Natl Cancer Inst 65*, 5 (1980), 1063–71.

[12] MORGAN, M. G. The neglected art of bounding analysis. *Environmental Science and Technology 35*, 7 (Apr.1 2001), 162A–164A.

[13] MORGAN, M. G., AND HENRION, M. *Uncertainty: A guide to dealing with uncertainty in quantitative risk and policy analysis*. Cambridge University Press, 1990.

[14] MORGAN, M. G., AND KEITH, D. Subjective judgments by climate experts. *Environmental Science and Technology 29*, 10 (Oct. 1995), 468A–476A.

[15] MORGAN, M. G., PITELKA, L. F., AND SHEVLIAKOVA, E. Elicitation of expser judgments of climate change impact on forest ecosystems. *Climatic Change 49* (2001), 279–307.

[16] ROCHA, L. M. Relative uncertainty and evidence sets: a constructivist framework. *International Journal of General Systems 26*, 1–2 (1997), 35–61.

[17] SAMET, J. M., BRENNER, D., BROOKS, A. L., ELLETT, W. H., GILBERT, E. S., GOODHEAD, D. T., HALL, E. J., HOPKE, P. K., KREWSKI, D., LUBIN, J. H., MCCLELLAN, R. O., AND ZIEMER, P. L. Health effects of exposure to radon, 1999.

[18] SCHOENBERG, J. B., STEMHAGEN, A., MASON, T. J., PATTERSON, J., BILL, J., AND ALTMAN, R. Occupation and lung cancer risk among new jersey white males. *Journal of the National Cancer Institute 79*, 1 (1987), 13–21.

[19] SHAFER, G. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton (NJ), 1976.

[20] SMETS, P. Belief functions and the transferable belief model, 2000.

[21] SMETS, P. Matrix calculus for belief functions, 2001.

[22] VINEIS, P., AND SIMONATO, L. Proportion of lung and bladder cancers in males resulting from occupation: a systematic approach. *Archives of Environmental Health 46*, 1 (1991), 6–15.

[23] WALLEY, P. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.

Minh Ha-Duong is Visiting Professor at the Engineering and Public Policy Department, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh PA 15213. E-mail: minh.ha.duong@cmu.edu

Elizabeth Casman is Research Engineer at the Engineering and Public Policy Department, Carnegie Mellon University.

M. Granger Morgan is Lord Chair Professor in Engineering; Professor and Department Head, Engineering and Public Policy; Professor, Electrical and Computer Engineering, and The H. John Heinz III School of Public Policy and Management, Carnegie Mellon University.