Risk Analysis, Vol. 30, No. 5, 2010
DOI: 10.1111/j.1539-6924.2010.01378.x

Probabilistic Inversion in Priority Setting of Emerging Zoonoses

Dorota Kurowicka,[1,*] Catalin Bucura,[1] Roger Cooke,[1,2] and Arie Havelaar[3]

This article presents methodology of applying probabilistic inversion in combination with expert judgment in priority setting problem. Experts rank scenarios according to severity. A linear multi-criteria analysis model underlying the expert preferences is posited. Using probabilistic inversion, a distribution over attribute weights is found that optimally reproduces the expert rankings. This model is validated in three ways. First, consistency of expert rankings is checked, second, a complete model fitted using all expert data is found to adequately reproduce observed expert rankings, and third, the model is fitted to subsets of the expert data and used to predict rankings in out-of-sample expert data.

KEY WORDS: Expert judgment; priority setting; probabilistic inversion; zoonoses

1. INTRODUCTION

Diseases that are naturally transmitted from warm-blooded animals to humans are called zoonoses. In Europe zoonoses originating from wildlife sources and transmitted by arthropods are considered to become increasingly important in the future. Climate and ecological changes may favor already existing arthropods to expand to other regions and thus new pathogens could be introduced in Europe. In 2007, the Dutch National Institute of Public Health and Environment (RIVM), together with several other research institutes, started a joint effort to monitor zoonoses and zoonotic agents in the Netherlands. One of the aims of this project is to build a model that allows prioritization of existing as well as emerging zoonoses in terms of their future threat. In this article we discuss analytic techniques used in building the model.

[1] Delft University of Technology, Delft, The Netherlands.
[2] Resources for the Future.
[3] RIVM, The Netherlands.
*Address correspondence to Dorota Kurowicka, Delft University of Technology, Mekelweg 4, 2628CD Delft, The Netherlands; d.kurowicka@tudelft.ne.

Priority setting is a multi-dimensional problem, in which technical information is often intertwined with value judgments. Several priority setting procedures have been used and described. Traditionally, a priority setting procedure entails asking a limited number of experts to reach consensus. An example of this approach in the domain of emerging zoonoses has been published in Reference 1.[(1)] Its method is relatively straightforward, but not very transparent and the repeatability is low. Currently, semi-quantive methods are frequently used in which criteria are divided into a limited number of classes (e.g., low, medium, and high). Criteria may also be scored on arbitrary scales (e.g., $0, 1, \ldots, 5$) and scores for all criteria are aggregated to produce an overall score. An example of this approach was published in Reference 2. Here, the transparency and the repeatability is improved, but the classes are chosen rather arbitrarly. Linear relations between the different classes of criteria or between criteria are often assumed but not founded in data. For this project, we aim at developing quantitative methods using natural numbers instead of arbitrary numbers. These methods are more transparant, objective, and can form the basis of future knowledge management systems.

The quantitative method is based on the well-established multi-criteria analysis (MCA) method. This method has been used in many decision-making contexts including animal health.[(3)] MCA offers methods and techniques to structure complex decision making. Generally, MCA consists of the following five phases:

(A) List and structure appropriate criteria (aspects of risk) to assess pathogens.
(B) Evaluate pathogens on the selected criteria.
(C) Determine the relative importance (weight) of each of criterion.
(D) Aggregate the scores and weights of the criteria into one overall value per pathogens.
(E) Perform sensitivity analysis.

This article will present a preliminary approach to phase A, and will focus on new mathematical techniques to address phase C. The results of phases B, D, and E will be reported elsewhere. The basic techniques applied here were also used in a study of marine ecosystem threats.[(4)]

Based on discussions between institutions participating in the project a prototype model with nine characteristics was chosen to capture the most relevant aspects of risk of a pathogen. These characteristics, which we call attributes, are given in Table I. Attributes should be measurable quantities that can be described on a monotonic scale. They have to be carefully and clearly defined. Each attribute has four or five levels, each associated with a value. We should be able to calculate or estimate these values for all pathogens. Lower numerical values of attributes correspond to the less threatening cases.

We stress that these attributes are only preliminary (full description will be published in elsewhere).

The attributes are expressed in different units. We therefore need to transform them to a monotonic scale on [0,1]. For attributes 2, 3, 5, and 8 we first used logarithmic transformation and then all attributes were scaled linearly such that the most threatening level took value 1.

The model has to combine values of all attributes for a pathogen in a numerical score that will be used

Table I. Attributes of the Model

	(1) Probability of Introduction in NL (in %)		(2) Speed of Spread Between Animals (in days)		(3) Economic Damage Animals (in Million of €)		(4) Probability of Transmission from Animal to Human (in Fraction of Human to Animal Contact Insuring Infection)		(5) Speed of Spread Between Humans (in Days)
1.	0	1.	10000 (no spread)	1.	5	1.	1:10,000	1.	10,000
2.	0.5	2.	30	2.	50	2.	1:1,000	2.	30
3.	5	3.	10	3.	500	3.	1:100	3.	10
4.	50	4.	3	4.	5,000	4.	1:10	4.	3
5.	100								
	(6) Severity of Illness, Morbidity (Calculated as a Product of Disability Weight and Duration)		**(7) Chance of Dying, Mortality in Human Population (in%)**		**(8) Economic Damage Humans (in Million of €)**		**(9) Risk Perception (in Number of Subjective Aspects that Apply)**		**Subjective Aspects Considered in 9)**
1.	0.02	1.	0	1.	5	1.	0		■ Involuntary exposure
2.	0.06	2.	0.5	2.	50	2.	2		■ Inequity (who profits)
3.	0.2	3.	5	3.	500	3.	4		■ Cannot be avoided through personal behavior
4.	0.6	4.	50	4.	5000	4.	6		■ Unknown and unnatural risk
		5.	100						■ Hidden, postponed, and irreversible damage
									■ Identification with victims (e.g., children)

Table II. Scenarios in Group 1, Showing Levels and Values

	Attributes																	
	1		2		3		4		5		6		7		8		9	
S_1	4	50	3	10	3	500	3	1/100	3	10	3	0.2	4	50	3	500	3	4
S_2	3	5	2	30	3	500	4	1/10	4	3	3	0.2	3	5	2	50	4	6
S_3	4	50	3	10	1	5	3	1/100	4	3	3	0.2	4	50	1	5	4	6
S_4	4	50	4	3	4	5,000	4	1/10	3	10	2	0.06	3	5	2	50	1	0
S_5	4	50	1	10,000	3	500	2	1/1,000	4	3	4	0.6	3	5	3	500	2	2
S_6	2	0.5	2	30	4	5,000	3	1/100	3	10	3	0.2	4	50	2	50	4	6
S_7	3	5	3	10	2	50	3	1/100	4	3	3	0.2	4	50	2	50	3	4

The values have physical dimension as given in Table I.

Table III. Experts' Assessments for Group 1

		EXPERTS										
SCENARIOS		1	2	3	4	5	6	7	8	9	10	11
1	QJ	JR	WL	ZC	ZC	WL	ZC	VG	VG	JR	ZC	ZC
2	VG	VG	PX	VG	JR	VG	PX	ZC	ZC	VG	PX	VG
3	GF	PX	ZC	PX	VG	ZC	JR	JR	JR	ZC	JR	PX
4	JR	QJ	VG	GF	WL	JR	GF	GF	WL	WL	VG	JR
5	ZC	ZC	JR	JR	PX	PX	VG	PX	PX	PX	WL	WL
6	WL	GF	QJ	WL	QJ	QJ	QJ	WL	QJ	QJ	GF	GF
7	PX	WL	GF	QJ	GF	GF	WL	QJ	GF	GF	QJ	QJ

to prioritize threat of all existing as well as emerging pathogens.

Severity scores are not observable, thus the idea emerged to elicit orderings of certain scenarios in terms of severity from experts. These scenarios do not describe any particular pathogen; they are random combinations of levels of each attribute. For our problem we have generated randomly 30 different scenarios reflecting hypothetical zoonoses. They were generated such that none of them "majorizes" the others, that is, there is no scenario whose attributes' values are greater or equal to the values of some other scenario. Randomly choosing scenarios avoids selection bias for the 30 scenarios.

The total set of scenarios was divided into six groups, each group consisting of seven scenarios. Scenarios in first groups are in general more severe then in last groups. In Table II, seven scenarios from group 1 are shown. Scenarios overlap between groups. In the first five groups the last two scenarios were repeated as first ones in the consecutive group. The last group contained the last four scenarios from group 5. This design was chosen for two reasons: first, the task of ordering 35 scenarios was judged to be cognitively too burdensome, and second, overlapping the scenarios enables consistency checks.

In the next section, we discuss elicitation of orderings for six groups of scenarios. Later these orderings will be used to find a model of scores for pathogens.

2. EXPERTS' DATA

Eleven experts participated in the elicitation. Prior to the elicitation the purpose of the study as well as the procedure was explained. Attributes and scenarios were explained. Each scenario was written on a card. Different groups of scenarios were distinguished by different colored cards. Experts were asked to arrange the cards from each of the six groups in increasing order of severity. Experts' assessments for the scenarios in the first group presented in Table II are shown in Table III. We see that the first scenario, called QJ, is ranked fourth by expert 1, sixth by expert 2, and seventh by expert 3. Overall, we see that experts considered QJ as very severe, much more severe than, for example, the seventh scenario, called PX.

Table IV. Summary of Experts' Assessments for Groups 1, 2, and 3

		Scores	1st	2nd	3rd	4th	5th	6th	7th
Group 1	S_1	68				1		6	4
	S_2	28	2	5	1	2	1		
	S_3	65				3		3	5
	S_4	34	2	1	4	2	2		
	S_5	23	5	2	3		1		
	S_6	50	2			3	2	2	2
	S_7	40	3	3			5		
Group 2	S_6	34	2	4	2		1	1	1
	S_7	40	1	3	1		6		
	S_8	32	3		5	2		1	
	S_9	62		1		1	1	5	3
	S_{10}	64	1			2		1	7
	S_{11}	42	1	1	1	5	2	1	
	S_{12}	34	3	2	2	1	1	2	
Group 3	S_{11}	53		1	1	2	3	3	1
	S_{12}	17	10						1
	S_{13}	55		2	1	1	2	1	4
	S_{14}	46			4	3	2	2	
	S_{15}	38	1	2	2	4	1	1	
	S_{16}	56		2	1	1	2		5
	S_{17}	43		4	2		1	4	

Table V. Summary of Experts' Assessments for Groups 4, 5, and 6

		Scores	1st	2nd	3rd	4th	5th	6th	7th
Group 4	S_{16}	33	3	2	3		1	2	
	S_{17}	40			5	5	1		
	S_{18}	49		3		1	4	2	1
	S_{19}	47	1		3	2	3		2
	S_{20}	67				2	2		7
	S_{21}	53	1	2			1	6	1
	S_{22}	20	6	4				1	
Group 5	S_{21}	60	2				2	1	6
	S_{22}	33	3	3	1	2		1	1
	S_{23}	55	1	1		1	1	6	1
	S_{24}	38		2	5	2	1	1	
	S_{25}	32	4	2	1	1	2		1
	S_{26}	45		3	2	2	1	1	2
	S_{27}	45	1		2	3	4	1	
Group 6	S_{24}	36	2	2	3	2		1	1
	S_{25}	35	4		2	1	3	1	
	S_{26}	49	1	3		2		1	4
	S_{27}	47	2	1		2	3	1	2
	S_{28}	45		2	2	3	2	1	1
	S_{29}	56			2	1	2	6	
	S_{30}	40	2	3	2		1		3

Experts' rankings for each group are summarized in Tables IV and V. The first scenario was ranked fourth by one expert (compare with Table III). Six experts ranked this scenario sixth and four experts ranked it seventh. The third column of these tables (denoted as "scores") shows the rank score of each scenario obtained by multiplying the number of experts who ranked scenario i as jth by its rank order, thus j. The rank score of the first scenario is 68 ($1 \times 4 + 6 \times 6 + 4 \times 7 = 68$).

Table VI. Values of S for Each Group

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
S	1,866	1,088	1,116	1356	680	340

Two different analyses of experts' orderings were performed. We first checked if experts were consistent when ordering repeated scenarios in different groups. Only one expert was consistent in all his/her assessments. One expert changed orderings of repeated scenarios in all groups. The others performed "reasonably," producing only one or two inconsistencies. Second, we have investigated if experts' orderings were given at random. The coefficient of concordance has been used for this purpose.[5,6] The hypothesis of giving orderings at random is tested by calculating the sum mean square difference (S) between the expert-wise average rank of a scenario and the overall average rank. For completeness we show in Table VI values of S for each group. Large S values are significant, and the hypothesis of random orderings would be rejected at the 5% level when $S = 630$. The hypothesis of random orderings is rejected for all groups except group 6. However, the agreement for group 5 is near the rejection threshold. We interpret this to mean that the ability to discriminate severity of scenarios drops off as the scenarios become less severe.

Because we cannot reject the hypothesis that orderings in group 6 are given at random, we have removed group 6 from our analysis.

3. MODEL

Given information provided by experts we must find a model that will satisfy experts' assessments. We opted for the simplest possible model satisfying requirements, the linear model. Failure of the linear model to provide an adequate fit would motivate a search for more complicated models. We define the score of each scenario as a linear combination of attributes' values:

$$S_i = \sum_{j=1}^{9} B_j X_{ij} \quad i = 1, \ldots, 30, \qquad (1)$$

where X_{ij} is a value of jth attribute in ith scenario. B_js are parameters of this model and they have to be found such that experts' constraints given in tables in Tables IV and V are satisfied. In particular, we regard each expert as having his/her own vector $(B_1 \ldots B_9)$ of coefficients, and hence we regard the population of experts as a distribution over possible values for $(B_1 \ldots B_9)$.

Probabilistic inversion based on samples reweighting was employed to find a distribution over $(B_1 \ldots B_9)$ that optimally reproduces the expert data. That is, by sampling from the optimal distribution over $(B_1 \ldots B_9)$, we should find that the percentage of experts who rank $S_i > S_j$ agrees with the data in Tables IV and V. In the next section, we explain this technique only on a simple example. For more information, including mathematical formulation, proofs, and applications, we refer the reader to References 7 and 8.

4. PROBABILISTIC INVERSION VIA SAMPLE REWEIGHTING

Suppose we have two independent, uniformly distributed random variables X_1 and X_2. Since they are independent, the probability of X_1 being bigger than X_2 is 0.5. The scatter plot with 10,000 samples of X_1 and X_2 in case of independence is shown in Fig. 1 (left). If we decide to impose the constraint $P(X_1 > X_2) = 0.8$ we must change the joint distribution of X_1 and X_2. For this purpose the probabilistic inversion based on samples reweighting can be used. We apply an iterative proportional fitting (IPF) algorithm to find new weights for samples such that after reweighting the constraint will be satisfied. Samples for which $X_1 > X_2$ will get a bigger weight. Instead of 1/10,000 their weight will be $0.8/(0.5 \times 10{,}000)$ and after resampling we will see more mass concentrated in the bottom right corner of the unit square; see Fig. 1 (middle). For only one constraint it is easy to see how weights have to change. In case more constraints have to be imposed IPF will change weights iteratively satisfying one constraint at a time. Csiszar[9] proves that IPF converges to a solution if the problem is feasible, that is, when a solution exists. Moreover, the limit is the minimally informative distribution satisfying the constraints, relative to the starting distribution. Fig. 1 (right) shows the scatter plot of the distribution obtained with IPF when two constraints were imposed, namely, $P(X_1 > X_2) = 0.8$ and $P(X_1 > 1 - X_2) = 0.8$. We see now more mass concentrated in the top right corner as well as the bottom right corner of the unit square.

5. ANALYSIS

We begin by analyzing each group separately and then combine groups 1–5. Combining all scenarios using all information provided by experts is ill-advised because we would disable out-of-sample validation, and because the number of constraints becomes intractable (each nonempty cell in columns 6 to 12 in Tables IV and V for groups 1–5 corresponds to a possible constraint). Hence our goal is to first see which constraints are sufficient to properly reconstruct the experts' preference percentages. We now discuss in detail group 2.

There are 35 nonempty cells in the last seven columns for the second group, which means that we have at most 35 constraints to impose on the joint distribution of scores. They are of the following type: for the sixth scenario we have the chance that S_6 is the lowest ranked within the second group is 2/11, the chance that S_6 is second from the bottom is 4/11, etc., and finally the chance that S_6 is ranked seventh (most severe) is 1/11.

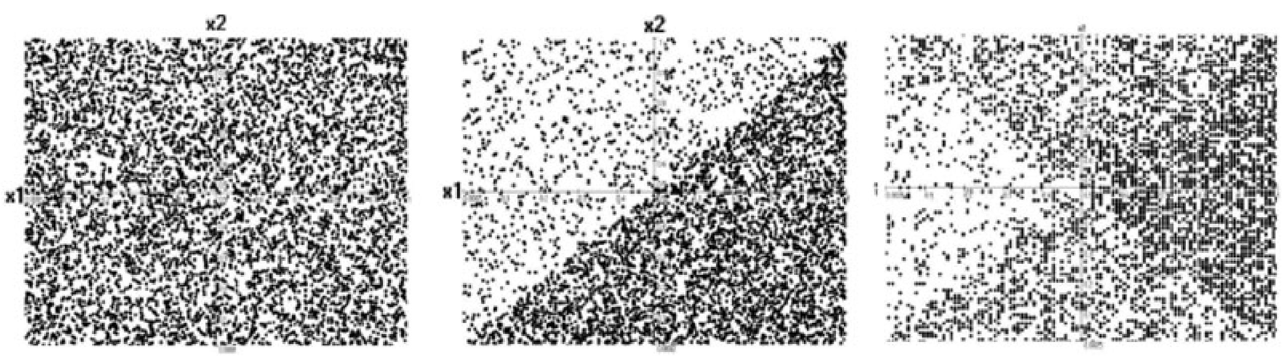

Fig. 1. Scatter plots in cases when $P(X_1 > X_2) = 0.5$ (left), $P(X_1 > X_2) = 0.8$ (middle), and when $P(X_1 > X_2) = 0.8$ and $P(X_1 > 1 - X_2) = 0.8$ (right).

Table VII. Mean and Variances of Coefficients of Equation (1) Under Variant I for Group 2

	Mean	Variance
B_1	0.7208	0.0643
B_2	0.5865	0.0761
B_3	0.2211	0.0637
B_4	0.2885	0.0549
B_5	0.2300	0.0443
B_6	0.3225	0.0453
B_7	0.5955	0.0626
B_8	0.5040	0.0971
B_9	0.6055	0.0472

We first impose all 35 constraints to see what the ordering of scenarios and their mean scores must be and then we will consider a few strategies with smaller numbers of constraints and compare their performance. To judge the performance of each variant we choose the following statistics:

1. **Error of scores** represents the square root of the mean square difference between mean scores obtained when imposing all constraints and scores calculated in a given variant. This statistic shows how well the variant can recover mean scores of the scenarios.
2. **Error of fitting** shows how well the model fits the summary data shown in Tables IV and V. It is calculated as the square root of the mean square difference between the imposed and obtained constraint probabilities.
3. **Error of validation** is computed as the square root of the mean square differences between the obtained probabilities while imposing all constraints and the probabilities recovered in a given variant.

After imposing all 35 constraints on the joint distribution of scores calculated with the linear model (Equation (1)), we discovered that the linear model fits the information provided by experts. The problem that IPF had to solve was feasible (differences between imposed and obtained probabilities were of order of 10^{-2}). Means and variances of B_js obtained in Variant I are shown in Table VII.

Means of scenario scores calculated with all constraints are shown in the second column in Table VIII under variant I.

Many variants with subsets of 35 constraints could be considered. We present the results for four such subsets with the short motivation behind each subset choice.

- *Variant I*. We denote as variant I the variant with all 35 constraints.
- *Variant II* contains all constraints except the ones concerning how many experts considered a given scenario as fourth in the ordering. Hence the constraints from the eighth column in Tables IV and V are not taken into account. For the sixth scenario the number of constraints in this variant is the same as in variant I as none of the experts considered this scenario as fourth.
- In *Variant III* only information on how many experts considered the rank of a given scenario as the lowest, the second lowest, the highest, and the second highest is included. For the sixth scenario, four constraints will be used. The probability that S_6 is the lowest is 2/11, second lowest 4/11, highest 1/11, and second highest also 1/11.
- *Variant IV* contains constraints concerned only with the lowest and the highest ranks. For the sixth scenario only the probability that S_6 is the lowest 2/11 and highest 1/11 must be imposed. We can call this variant "smallest and highest."
- *Variant V* contains only those constraints on which at least three experts agreed that a given scenario should be ranked at a given position.

Table VIII. Scores and Orderings of Scenarios in Group 2 for Different Variants

	Variants				
Group 2	I	II	III	IV	V
Ordered Scenarios & Scores	$S_{10} = 1.529$	$S_{10} = 1.541$	$S_{10} = 1.524$	$S_{10} = 1.556$	$S_9 = 1.474$
	$S_9 = 1.503$	$S_9 = 1.538$	$S_9 = 1.523$	$S_9 = 1.536$	$S_{10} = 1.464$
	$S_{12} = 1.435$	$S_{11} = 1.436$	$S_{11} = 1.409$	$S_{11} = 1.473$	$S_{11} = 1.409$
	$S_{11} = 1.434$	$S_6 = 1.370$	$S_6 = 1.356$	$S_6 = 1.448$	$S_{12} = 1.397$
	$S_6 = 1.390$	$S_{12} = 1.345$	$S_{12} = 1.350$	$S_{12} = 1.391$	$S_6 = 1.381$
	$S_8 = 1.366$	$S_7 = 1.318$	$S_8 = 1.312$	$S_7 = 1.372$	$S_7 = 1.314$
	$S_7 = 1.326$	$S_8 = 1.298$	$S_7 = 1.304$	$S_8 = 1.301$	$S_8 = 1.286$

Table IX. Errors in Group 2

Group 2 Variant	Number of Constraints	Error of Scores	Error of Fitting	Error of Validation
I	35	0	0.2226	0
II	30	0.4848	0.9154	0.9298
III	20	0.4950	2.0323	2.1003
IV	9	0.5581	3.5302	3.6110
V	10	0.5371	1.6301	1.6914

In this variant, for scenario S_6 only one constraint will be imposed, namely, the probability that S_6 is second lowest is 4/11. We will refer to this variant as "at least 3."

Variants II, III, and IV are easily motivated as we are mainly concerned with separating the most from the least threatening scenarios. As long as the most severe and the least severe scenarios are recognized it is not so bad if orderings of scenarios ranked in the middle of the group are changed in the model. Variant V, "at least 3," is of different type. We adopted here an idea that only strong signals coming from experts have to be seriously taken into account. All other constraints can be treated as noise.

The analysis was done using the same 100,000 initial samples in each variant. In Table VIII mean scores are shown. We can see that scores don't change much between different variants. Moreover, orderings of scenarios remain similar recovering scenarios 9 and 10 as the most severe and scenarios 7 and 8 as the least severe in this group.

To decide which variant is the most appropriate for the further analysis we compare three statistics introduced above: Error of scores, Error of fit, and Error of validation. They are presented together with number of constraints in Table IX. Variant V, "at least 3," has the best performance in terms of all three statistics. Moreover, it requires only 10 constraints to be imposed on the joint distribution of scores. This is one more than in variant IV, "lowest and highest," but the performance of variant IV in terms of fitting and validation error is much worse.

Similar conclusions have been drawn for other groups. We noticed that for all groups (except the sixth one, which was removed from the analysis) the linear model for scores was appropriate. We have observed that when fitted separately quite different distributions of parameters of the linear model were obtained. This could suggest that different models should be used for more severe and less severe scenarios. The best performing variant was variant V, "at least 3." Hence, we used this variant to combine all scenarios together. The relatively low validation errors for variants II–V indicate that a fitting based on a relatively small subset of constraints enables reasonable predictions of the expert response rates for the out-of-sample constraints, and provides out-of-sample validation for the linear model in general.

Table X. Mean and Variances of Coefficients of Equation (1) Under Variant V for Groups I–V

	Mean	Variance
B_1	0.5570	0.0930
B_2	0.2210	0.0295
B_3	0.2988	0.0474
B_4	0.2965	0.0608
B_5	0.3394	0.0608
B_6	0.1521	0.0261
B_7	0.5159	0.1027
B_8	0.6207	0.0883
B_9	0.2004	0.0546

To build one model for all 27 scenarios from groups 1, 2, 3, 4, and 5 using variant V we had to impose 53 constraints on the joint distribution of their scores. Because of this high number of constraints and attendant computational restrictions, some infeasibility was observed. The differences in imposed and obtained probabilities were sometimes of order of 0.2. In Table X means and standard deviations of parameters of the linear model of scores are presented.

The best fitting joint distribution over $(B_1 \ldots B_9)$ correlates these values. The correlation matrix is shown in Table XI. Several coefficients exhibit moderate correlations. This table indicates, for example, that coefficients B_3 and B_2 are correlated at 0.4664. The model thus predicts that high valuations of B_3 tend to appear with high values of B_2. If a new scenario with new attribute scores is constructed, the valuation of this scenario by a randomly drawn expert from the modeled population of experts would be given by using the mean values in Table X. This of course is the great advantage of the linear model (Equation (1)). Had the linear model proved inadequate, then we should have to contemplate adding interaction and/or higher order terms.

We can notice that the probability of introduction (attribute 1) and the mortality (attribute 7), as well as economic damage related to human population (attribute 8), described in Table I, will have the

Table XI. Correlation Matrix

	B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8	B_9
B_1	1								
B_2	−0.1806	1							
B_3	−0.3992	0.4664	1						
B_4	−0.0172	−0.1096	−0.0915	1					
B_5	−0.0730	−0.2143	0.0387	−0.1256	1				
B_6	−0.2817	−0.2089	0.0256	−0.2352	−0.0525	1			
B_7	−0.1100	−0.3034	−0.4929	0.0481	0.2031	0.2415	1		
B_8	−0.1316	−0.4633	−0.2265	0.1229	0.4128	0.1135	0.3864	1	
B_9	0.2597	0.2486	−0.0322	−0.0118	−0.3785	−0.0201	0.0849	−0.0721	1

highest weight in the severity score of a pathogen. It is a bit surprising that economic damage has such significance in determining the score, but this might be explained by relatively high positive correlations of this attribute with mortality (attribute 7) and speed of spread between humans (attribute 5).

6. DISCUSSION AND CONCLUSIONS

This study deploys a new method for modeling expert stakeholder preferences based on probabilistic inversion of expert ordinal rankings. This offers several advantages above traditional MCA approaches. First and foremost, experts are required only to rank observable scenarios with respect to severity. They do not rank or evaluate importance of attributes directly. Since attribute weights are multiplied with scores on physically measurable attributes, direct valuation of attribute weights would require experts to take account of the physical dimensions of all the attributes. Thus, if speed of spread among animals (attribute 2) were measured in weeks instead of days, the coefficient for attribute 2 would have to be adjusted accordingly. Second, the linear model (Equation (1)) can be validated by predicting percentage expert rankings for out-of-sample constraints. Such validation is not readily available for traditional MCA methods. Finally, the probabilistic inversion yields more information than those traditional methods. The best fitting distribution over $(B_1 \ldots B_9)$ yields not only mean values for the attribute weights, but also gives variances and correlations. These advantages come at a price. Probabilistic inversion is a sophisticated mathematical tool. In spite of compliant software to support this task, it does place demands on the analyst to explain the results to lay problem owners. Finally, the probabilistic inversion with the simple linear model (Equation (1)) may not work. There may be no joint distribution over $(B_1 \ldots B_9)$ that adequately reproduces the expert rankings. If this were to arise, the analyst would have to contemplate more complex models or may have to search for additional "hidden" attributes. This last point is inherent in any method that strives for external validation: the validation may not be forthcoming. Fortunately, this was not the case in the present study.

ACKNOWLEDGMENTS

We gratefully acknowledge Marieta Braks for participating in overall design of the system and managing the process, Juanita Haagsma for assistance in the panel sessions, Monique Mourits for help in designing the economic attributes, and the EmZoo consortium. This project was funded by the Ministry of LNV.

REFERENCES

1. Valenciano M. Définition des priorités dans le domaine des zoonoses non-alimentaires 2000–2001. Paris, France: Institut de Veille Sanitaire, 2001.
2. Anonymous. Prioritization approach project—Impact on public health workshop. A similar approach for wildlife pathogens, including some weights is in McKenzie J, Simpson H, Langstaff I. Development of methodology to prioritise wildlife pathogens for surveillance. Prev Vet Med 2007; 81(1–3)194–210. Available at http://www.defra.gov.uk/animalh/diseases/vetsurveillance/pdf/workshop_publichealth.pdf. Accessed on June 20, 2007.
3. Mourits M, Oude Lansink AGJM. Multi-criteria decision making to evaluate quarantine disease control strategies. Pp. 131–144 in Oude Lansink AGJM (ed). New Approaches to the Economics of Plant Health, The Netherlands: Waqeningen UR Frontis Series, 2007.
4. Neslo R, Micheli F, Kappel CV, Selkoe KA, Halpern BS, Cooke RM. Modeling stakeholder preferences with probabilistic inversion: Application to prioritizing marine ecosystem vulnerabilities. Pp. 248–271 in Linkov I, Ferguson E, Magar V (eds). Real Time and Deliberative Decision Making:

Application to Risk Assessment for Non-Chemical Stressors. Amsterdam: Springer.
5. Cooke RM. Experts in Uncertainty. Oxford, UK: Oxford University Press, 1991.
6. Siegel S. Nonparametric Statistics. New York: McGraw-Hill, 1958.
7. Du C, Kurowicka D, Cooke RM. Techniques for generic probabilistic inversion. Computational Statistics and Data Analysis, 2006; 50:1164–1187.
8. Kurowicka D, Cooke RM. Uncertainty Analysis with High Dimensional Dependence Modelling. New York: Wiley, 2006.
9. Csiszar I. I-divergence geometry of probability distributions and minimization problems. Annals of Probability, 1975; 3:146–158.