Comparison of sensitivity analysis techniques for a simulation-based risk assessment model

Heller J.[1], Innocent G.T.[2], Kelly L.[3,4], Reid S.W.J.[1] & Mellor D.J.[1]
[1] Boyd Orr Centre for Population and Ecosystem Health, Institute for Comparative Medicine, Faculty of Veterinary Medicine, University of Glasgow, Bearsden Road, Glasgow, G61 1QH, UK.
[2] Biomathematics and Statistics Scotland (BioSS), Edinburgh, UK.
[3] Department of Statistics and Modelling Science, University of Strathclyde, UK.
[4] Centre for Epidemiology and Risk Analysis, Veterinary Laboratories Agency, UK

Contact: j.heller@vet.gla.ac.uk

ABSTRACT

This paper presents the results of a sensitivity analysis using differing techniques for a simulation-based stochastic model assessing the probability of acquisition of meticillin-resistant *Staphylococcus aureus* (MRSA) in pet dogs. Three separate sensitivity analyses were undertaken. The simulation model to which the sensitivity analyses were applied was represented by seven pathways by which MRSA could be acquired, stratified by attendance at a veterinary clinic and run over a 24 hour period. Two sub-steps were modelled within each pathway: 1) exposure to a source of MRSA and 2) transmission, given exposure has occurred. Therefore, input factors could be considered at two levels: the initial input (raw data) level, or the intermediate level, representing the simulated exposure and transmission probabilities. The sensitivity analyses undertaken were 1) logistic regression (LR) modelling using intermediate transmission parameters, with and without consideration of interaction terms, 2) one-at-a-time (OAT) parameter variation for factors at the initial input level, and 3) a fractional factorial two level Plackett-Burman (P-B) design with foldover for factors at the initial input level. The results of the sensitivity analyses were complicated and ambiguous and the overall results of the three analyses were found to be markedly different. The resulting subjectivity of outputs is in contrast to the aim of the sensitivity analysis procedure and results of any sensitivity analysis should be interpreted with respect to the method used and the original model that it represents.

KEYWORDS
Sensitivity analysis, MRSA, fractional factorial design.

INTRODUCTION

Sensitivity analysis may be defined as the study of how the variation in the output of a mathematical model can be apportioned, qualitatively or quantitatively, to different sources of variation in the input of a model (Saltelli, 2000). The use of sensitivity analysis is considered to be an integral and essential component of the modelling process, the results of which may be used for numerous purposes including model validation, optimisation and calibration (Saltelli, 2000, Trocine and Malone, 2000, Borgonovo, 2008). In risk analysis, sensitivity analysis is primarily used to identify risk-governing parameters, important for both the implementation of mitigation strategies and direction of future research in the form of data collection and/or surveillance (Saltelli, 2000, Frey and Patil, 2002, Saltelli, 2002). A thorough and extensive sensitivity analysis allows the identification of influential input variables and priority data gaps and is particularly important in data sparse areas, where the outputs from the sensitivity analysis may represent the most influential output of a model. However, the term sensitivity analysis encompasses many different methods. This research aimed to compare and compile the results of different methods of sensitivity analysis when applied to the same scenario: a simulation-based risk assessment model in the data sparse area of acquisition of MRSA in dogs.

MATERIALS AND METHODS

The simulation model to which the sensitivity analyses were applied was developed to simulate the proportion of dogs that would acquire MRSA, as carriage, colonisation or infection over a 24 hour period (Heller, 2009). The model structure was based on a conceptual model that outlined the likely pathways of acquisition of MRSA for a single dog over a 24 hour time period. Briefly, seven pathways for acquisition were specified and stratified to represent human family member, non-family member or veterinary worker; animal, limited to dogs in the community or at veterinary clinics; and environmental sources of MRSA that could be accessed through community and veterinary hospital routes. The seven separate pathways for acquisition that were specified were considered to be non-sequential, were not mutually-exclusive and were stratified by attendance at a veterinary clinic during the given 24 hours. Two sub-steps were modelled within each pathway: 1) exposure to a source of MRSA and 2) transmission, given exposure has occurred. Therefore, input factors could be considered at two levels: the initial input (raw data) level, or the intermediate level, representing the simulated exposure and transmission probabilities. The sensitivity analyses undertaken were 1) stepwise logistic regression (LR) modelling including consideration of interaction terms, with outcome specified as MRSA status and potential explanatory variables as intermediate transmission variables from the simulation model, 2) OAT parameter variation for factors at the initial input level, and 3) a fractional factorial two level P-B design with foldover, also for factors at the initial input level. Two analyses were implemented for each method to account for all animals (5, 19 and 19 factors considered respectively) and those that visited a veterinary clinic (7, 28 and 28 variables respectively) in the period of time under consideration. This allowed for assessment of within-veterinary factors and avoided the inclusion of factors that were mutually exclusive.

RESULTS

The results of the three sensitivity analyses were varied and resulted in markedly different rankings of importance of input variables (Table 1). While the effect of the environment dominated the results of the LR models, significant interactions were also found between the probability of transmission of MRSA from the environment (home and veterinary) and all alternative sources. For the OAT analyses, the effect on the outcome was greatest for factors representing transmission from family members, non-family members, home environment and dogs within the community, in decreasing order. The P-B analyses found the five most influential variables to be the probability that a community dog is MRSA positive, the number of contacts with a non family member, the probability that the general human population is MRSA positive, the probability of transmission from a community dog and the number of contacts with family members, in decreasing order. Similar to the LR models, the results of the OAT and P-B analyses showed transmission from the veterinary environment to be the most influential factor for dogs that attended a veterinary clinic but the effect of all other within veterinary clinic factors varied.

DISCUSSION

The results of these analyses showed that different outputs and inference may result depending on the choice and implementation of sensitivity analysis technique. The stochastic model to which the sensitivity analyses were applied in this study was non-linear with input factors that were associated with numerous explicit and non-explicit dependencies, and uncertainties that varied in magnitude. As a result of these model-dependent factors, the outcomes of the three sensitivity analysis methods varied. The use of LR modelling for sensitivity analysis presents advantages over local sensitivity analysis models which include; 1) the ability to evaluate the sensitivity of individual model inputs while taking into account the simultaneous impact of other model inputs on the result (Helton and Davis, 2000, Frey and Patil, 2002) and 2) the ability to assess the input variables over a wide range, particularly when continuous variables are included. However, the reliance on a functional form of the relationship between the input and output variables, and the requirement for the key assumptions of regression to be met, represent potential limitations of the application of this method (Frey and Patil, 2002). The results of the OAT analyses are representative of a largely deterministic implementation of the simulation model and, while the weaknesses of an analysis such as this, including the inability to consider interactions, unsuitability to non-linear models and inability to resolve uncertainties of different orders of magnitude, have long been recognised (Henderson-Sellers and Henderson-Sellers, 1996, Campolongo et al., 2000a, Saltelli, 2000), these techniques continue to be commonly applied. The final technique, a P-B fractional factorial design reflects a method that allows for main effects and also potentially for factor interactions by varying the levels of multiple factors simultaneously (Campolongo and Saltelli, 2000, Campolongo et al., 2000b).

The overall results of the three analyses applied in this study were found to be markedly different. While the results of the OAT analysis may be described as more intuitive than the results of the P-B analysis, the knowledge that the stochastic model is non-linear and that the uncertainties associated with each factor are not always within the same order of magnitude, reduces confidence in the OAT results. Conversely, the P-B results are less intuitive, but it is known that the P-B design is more robust to the non-linearity and variation in input factor magnitudes that exist within the specified stochastic model (Beres and Hawkins, 2001). The results of the LR models were difficult to interpret and differentiate between but allowed consideration of putative interactions between input factors.

Table 1 Combined table of ranks of results for the LR, OAT and P-B sensitivity analyses for the simulation model for acquisition of MRSA in dogs.

Variables	All dogs			Dogs that attend veterinary clinics		
	OAT rank^	P-B rank§	LR rank^^	OAT rank^	P-B rank§	LR rank^^
P(dog attends vet in 24 hours)			#	NA	NA	NA
P(health care worker is MRSA positive)	10		*			*
P(veterinary worker is MRSA positive)			*			*
P(other high risk group is MRSA positive)			*			*
P(general population is MRSA positive)	8	3	*			*
P(person MRSA positive if lives with MRSA positive person)						
P(transmission from MRSA positive family member per contact)	1	6	2	5		=4
Number of contacts with family member per 24h	7	5				
P(transmission from MRSA positive non-family member per contact)	2	9	4	6		=5
Number of contacts with non-family member per 24h	9	2				
P(community dog is MRSA positive)		1			7	
P(transmission from MRSA positive community dog per contact)	4	4	3			=5
Number of contacts with community dog per 24h		7				
P(house contaminated with MRSA | at least one positive family member)						
P(house contaminated with MRSA | no positive family members)	6	10				
P(environmental site contaminated | at least one positive family member)			1**			1**
P(environmental site contaminated | no positive family members)					4	
P(transmission from MRSA positive home environment per contact)	3	8		7		
Number of contacts with home environment per 24h	5					
P(dog at vet is MRSA positive)	NA	NA	NA	8	5	
P(transmission from MRSA positive dog at veterinary clinic per contact)	NA	NA	NA	4	9	3
Number of contacts with a dog at a veterinary clinic per 24h	NA	NA	NA		3	
P(veterinary environment contaminated)	NA	NA	NA	2	2	
P(transmission from MRSA positive veterinary environment per contact)	NA	NA	NA	1	1	2**
Number of veterinary environment contacts per 24h	NA	NA	NA	9		
P(small animal veterinary worker is MRSA positive)	NA	NA	NA		6	
P(transmission from MRSA positive veterinary staff member per contact)	NA	NA	NA	3	10	=4
Number of veterinary staff contacts per 24h	NA	NA	NA		8	

P(..) = Probability; ^ Ranking of the first 10 factors that have a positive effect on the baseline ≥ a factor of two; § Ranking of the first 10 factors that have a positive main effect; ; ^^ Ranked based on odds ratios of main effects. Boxes denote all input variables that are included in the estimate of composite input factor;
Not included in ranking as dichotomous variable; * Also included in non family member composite factor; ** When P(transmission) from all other routes = 0

CONCLUSION

In conclusion, this study found that the chosen technique and resolution of input variables is likely to have a marked effect on the output of any sensitivity analysis undertaken. The method chosen, number of factors included in the analysis and the point in the model from which these factors were obtained, are likely to result in variation in the results of any sensitivity analysis that is undertaken. The resulting subjectivity of outputs is in contrast to the aim of the sensitivity analysis procedure. In short, the results of all sensitivity analyses should be interpreted with respect to the method used and the original model that it represents.

REFERENCES

Beres, D. L. & Hawkins, D. M. (2001) Plackett-Burman techniques for sensitivity analysis of many-parametered models. *Ecological Modelling,* 141**,** 171-183.

Borgonovo, E. (2008) Sensitivity analysis of model output with input constraints: a generalized rationale for local methods. *Risk Anal,* 28**,** 667-80.

Campolongo, F., Kleijnen, J. P. C. & Andres, T. (2000a) Screening methods. IN SALTELLI, A., CHAN, K. & SCOTT, E. M. (Eds.) *Sensitivity Analysis.* Chichester, John Wiley & Sons Ltd.

Campolongo, F. & Saltelli, A. (2000) Design of Experniments. IN SALTELLI, A., CHAN, K. & SCOTT, E. M. (Eds.) *Sensitivity Analysis.* Chichester, John Wiley and Sons Ltd.

Campolongo, F., Saltelli, A., Sorensen, T. & Tarantola, S. (2000b) Hitchhiker's Guide to Sensitivity Analysis. IN SALTELLI, A., CHAN, K. & SCOTT, E. M. (Eds.) *Sensitivity Analysis.* Chichester, John Wiley and sons Ltd.

Frey, H. C. & Patil, S. R. (2002) Identification and review of sensitivity analysis methods. *Risk Anal,* 22**,** 553-78.
Heller, J. (2009) A risk assessment approach to meticillin-resistant *Staphylococcus aureus* in pet dogs. *Department of Animal Production and Public Health.* Glasgow, University of Glasgow.
Helton, J. C. & Davis, F. J. (2000) Sampling-Based Methods IN SALTELLI, A., CHAN, K. & SCOTT, E. M. (Eds.) *Sensitivity Analysis.* Chichester, John Wiley and sons Ltd.
Henderson-Sellers, B. & Henderson-Sellers, A. (1996) Sensitivity evaluation of environmental models using fractional factorial experimentation. *Ecological Modelling,* 86**,** 291-295.
Saltelli, A. (2000) What is Sensitivity Analysis? IN SALTELLI, A., CHAN, K. & SCOTT, E. M. (Eds.) *Sensitivity Analysis.* Chichester, John Wiley & Sons.
Saltelli, A. (2002) Sensitivity analysis for importance assessment. *Risk Anal,* 22**,** 579-90.
Trocine, L. & Malone, L. C. (2000) Finding important independent variables through screening designs: A comparison of methods. *Proceedings of the 2000 Winter Simulation Conference, Vols 1 and 2***,** 749-754
2114.

ACKNOWLEDGEMENTS

This work was funded by DEFRA and SFC on the VTRI 0101 project.