1999

A Bayesian Approach to Sensitivity Analysis

Felli, James C.; Hazen, Gordon B.

Health Economics Letter

Health Economics, Volume 8, pp. 263-268, 1999.
http://hdl.handle.net/10945/44157

HEALTH ECONOMICS
Health Econ. **8**: 263–268 (1999)

HEALTH ECONOMICS LETTER

A BAYESIAN APPROACH TO SENSITIVITY ANALYSIS

JAMES C. FELLI[a,*] AND GORDON B. HAZEN[b]
[a] *Naval Postgraduate School, Monterey, CA, USA*
[b] *Northwestern University, Evanston, IL, USA*

SUMMARY

Sensitivity analysis has traditionally been applied to decision models to quantify the stability of a preferred alternative to parametric variation. In the health literature, sensitivity measures have traditionally been based upon distance metrics, payoff variations, and probability measures. We advocate a new approach based on information value and argue that such an approach is better suited to address the decision-maker's real concerns. We provide an example comparing conventional sensitivity analysis to one based on information value. This article is a US government work and is in the public domain in the United States.

KEY WORDS — Bayesian decision theory; the value of information; economics of information; statistical methods

INTRODUCTION

Sensitivity analysis (SA) refers broadly to any analytic method designed to quantify the impact of parametric variation on model output. For health-related decision models, measures of sensitivity based upon threshold proximity, range of value, and probability of decision change are well-established but suffer from severe limitations. We take the Bayesian position that all uncertainty should be quantified by probability distributions, *including* the uncertainty about parameter values which motivates SA, and propose a SA method based on information value which not only surmounts the limitations of conventional SA, but offers several advantages. We illustrate this approach using a published medical decision analysis and compare the results with those of an established SA method.

SENSITIVITY ANALYSIS

Sensitivity analysis *methods* have been partitioned into four categories in the health economics literature [1]: simple SA, threshold analysis, analysis of extremes, and probabilistic SA. The corresponding *measures* of the degree of sensitivity have typically been based on the distance between parameter point estimates and points of decision change (simple SA, threshold analysis), the range of model output swings consequent to parametric variation (analysis of extremes), and the relative frequency of an alternative's optimality (probabilistic SA). For context, suppose a decision-maker (DM) must choose among alternatives a having expected payoffs $E[V_a]$, and suppose alternative a^* is optimal, that is $E[V_{a^*}] = \max_a E[V_a]$. Let ξ be a problem parameter whose precise value is uncertain, and let $a^*(\xi)$ be the

* Correspondence to: DRMI (Code 64 FL), Naval Postgraduate School, 1522 Cunningham Road, Monterey, CA 93943-5201, USA. E-mail: jcfelli@nps.navy.mil

Contract/grant sponsor: Decision, Risk and Management Science Program at the National Science Foundation

Received 1 August 1998
Accepted 7 December 1998

optimal alternative given ξ, that is, $E[V_{a^*(\xi)}|\xi] = \max_a E[V_a|\xi]$.

Threshold proximity measures focus on proximity to a threshold in parameter space. For example, if $a^*(\xi) = a^*$ only when $\xi \leq \xi^*$ then ξ^* is a decision threshold for ξ. If the DM believes it sufficiently likely that $\xi > \xi^*$, he would label a^* sensitive to ξ. Graphical analyses give the same information, but illustrate the situation pictorially (see Plante *et al.* [2] for an example).

Range-of-value measures are commonly presented as tornado diagrams, using horizontal bars to illustrate the range of payoff values expected as parameters vary. Any bar extending beyond a predefined payoff threshold indicates potential sensitivity. A bar extending sufficiently far beyond the threshold raises concern about problem sensitivity to that parameter (see Dippel *et al.* [3] for an example).

Probabilistic sensitivity measures characterize sensitivity in terms of the probability of decision change. Because of the uncertainty in the parameter ξ, the optimal decision $a^*(\xi)$ conditioned on ξ is also uncertain. In probabilistic sensitivity analysis one calculates $P[a^*(\xi) = a]$ for all possible alternatives a. If $P[a^*(\xi) = a^*]$ is large, the DM would not consider the problem sensitive to ξ, regardless of how close ξ is to its threshold (see Doubilet *et al.* [4] for an example). Probabilistic approaches are consistent with the Bayesian perspective in that they require the DM to provide a distribution for ξ.

Limitations to traditional sensitivity measures include a lack of a formal definitions of 'sensitivity' [5,6] and difficulties in defining distance metrics [7,8], especially in the multiparametric case when ξ is a vector of parameters [9,10]. Evidence also suggests that conventional SA systematically overestimates problem sensitivity [11,12].

THE EXPECTED VALUE OF PERFECT INFORMATION

The expected value of perfect information on ξ, which we denote EVPI(ξ), is the difference between optimal expected payoffs with and without perfect knowledge of ξ prior to the time of decision [13,14]. That is, $\text{EVPI}(\xi) = E_\xi[E[V_{a^*(\xi)}|\xi]] - E[V_{a^*}]$. Because EVPI($\xi$) represents the average benefit consequent to resolving all uncertainty surrounding ξ prior to the point of decision, we contend that a decision problem should only be considered sensitive to ξ when EVPI(ξ) exceeds some minimally significant amount defined by the DM [15–18].

There are several advantages to EVPI-based sensitivity measures [12]. Because EVPI calculation requires parameter distributions, the DM's beliefs about parameter behaviour must be formalized sooner rather than later in the decision process. Also, EVPI calculation remains tractable when ξ is a vector. While these two advantages are also intrinsic to probabilistic SA, an EVPI analysis offers two additional benefits. First, EVPI(ξ) can be expressed as the product of the probability of a change in the optimal alternative due to variation in ξ and the average foregone payoff given such a change, encapsulating likelihood and value issues in a single measure. Second, the DM's ultimate assessment of sensitivity is based on values in natural units (e.g., dollars, quality adjusted life years) representing expected marginal benefits commensurate with complete resolution of parametric uncertainty.

AN EXAMPLE OF EVPI-BASED SA

A decision analysis of treatment options for pyriform sinus cancer [2] is provided in Figure 1. In addition to probabilities (e.g., d_r and m_r), the authors employed quality adjustment factors and tolls to weeks of survival awarded. Quality adjustment factors were assigned for surgery (q_s) and radiation (q_r); for mixed treatment strategies, the product of the pure treatment quality adjustment factors was used. Tolls represented direct reductions in quality adjusted life duration for surgery (T_s) and morbidity following surgery (T_{sm}) and radiation (T_{rm}).

Baseline analysis shows Surgery Then Post-Operative Irradiation to be the optimal alternative with an expected value of 108 additional quality-adjusted weeks of survival (QAWS). The authors performed traditional one-way, graphical SA on the probability of surgical mortality (d_s), the probability of disease-free survival with surgery (p_s), the fold-increase in disease-free survival for combined surgery/irradiation therapies (f_i), and the short-term morbidity of surgery (T_s). They also provided three graphical, two-way SAs: p_s versus d_s; f_i versus q_s (quality of life after surgery); q_s versus q_r (quality of life after radiation ther-

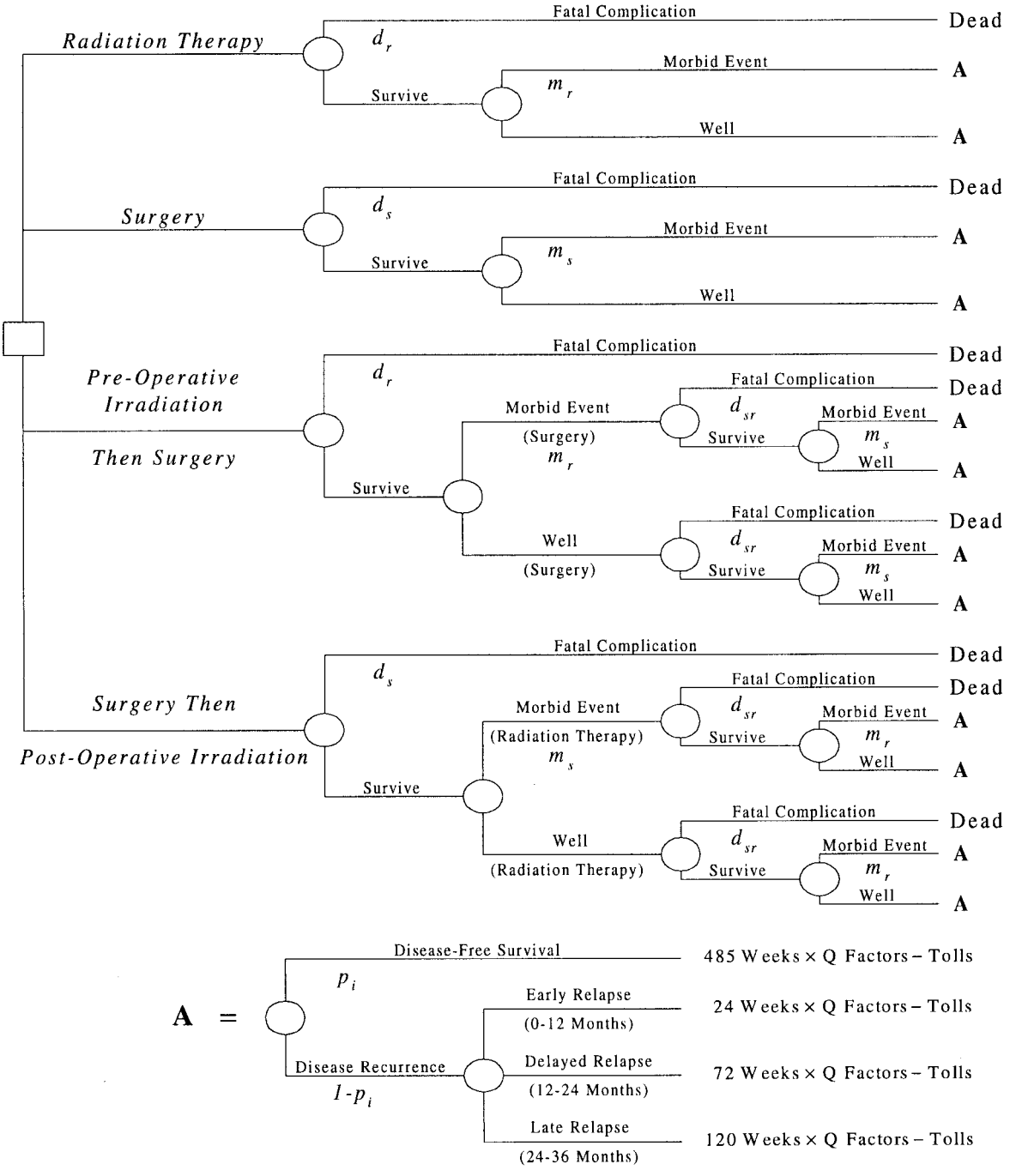

Figure 1. The decision tree for selecting a treatment for pyriform sinus cancer (Plante *et al.* [2]). In the outcome subtree, the probability p_i is the probability of disease-free survival, where i designates a pure treatment alternative: surgery ($i = s$) or radiation therapy ($i = r$). If the treatment strategy includes both irradiation and surgery then $p_i = f_i \cdot p_s$, where f_i is the 'fold-increase' in disease-free survival.

apy). They concluded the optimal alternative to be quite sensitive to f_i, somewhat sensitive to d_s, p_s, q_r, and q_s, and insensitive to T_s.

To compare these sensitivity conclusions with an EVPI analysis, we assigned parameter distributions (Table 1) based on the information the authors provided (including statements of plausible parameter ranges, ranges of values used in SA, and informal commentary) and used Monte Carlo simulation to estimate $\text{EVPI}(\xi) = E_\xi[E[V_{a^*(\xi)}|\xi]] - E[V_{a^*}]$ for parameter sets ξ (for a detailed description of the simulation procedure, see Felli and Hazen [11]). We assumed parametric independence. (This assumption does not affect our comparisons between one-way SA and one-way EVPI, and is unlikely, we believe, to affect comparisons in the multiway case—see Felli and Hazen [12].) Table 2 presents the EVPIs for all one- and two-element parameter sets ξ, as well for the entire parameter set. The parameters the authors judged sensitive according to their SAs are shaded.

To what parameter combinations does EVPI indicate sensitivity? The total-parameter EVPI of 6.066 QAWS constitutes 5.62% of the base-optimal 108 QAWS. This seems significant, so the problem appears sensitive to its parameter set as a whole. The parameter sets $\{p_r, q_r\}$ and $\{p_r, d_s\}$ possess EVPIs constituting 3.29% and 2.95% of the base optimal 108 QAWS. It is not clear whether improvements of this magnitude are significant, so it is questionable whether the problem is sensitive to these parameter sets. All remaining parameter sets have EVPIs less than 2.7%. Sensitivity in these cases seems marginal at best. Granted, these judgements are but reasonable conjectures on our part—the final word would rest with the DM. Even so, the parameters the authors judged sensitive via conventional SA (shaded in Table 2) are clearly inconsistent with our EVPI results. The single parameter with the highest EVPI (p_r) was not deemed sensitive, and all of the five parameters which were have information values less than 1.2% of the optimal 108 QAWS. This low rate of 'missed sensitivity' and high rate of 'false sensitivity' relative to an EVPI analysis is consistent with our general findings that conventional SAs typically overstate problem sensitivity to input parameters [12]. In this example, we compared threshold proximity based SA to an EVPI analysis. For a contrast between probabilistic SA and EVPI, we refer the reader to Felli and Hazen [11].

Table 1. Parameter distributions based on the information including statements of plausible parameter ranges, ranges of values used in SA, and informal commentary)

Model parameter examined	Symbol	Min	Lower	Base	Upper	Max	Distribution used
Surgery Mortality	d_s	0	0.01	0.05	0.1	1	Piecewise Linear (·, 0.05)
Combined Surg/Rad Mortality	d_{sr}	0	0.03	0.05	0.14	1	Piecewise Linear (·, 0.05)
Disease-Free Survival, Radiation	P_r	0	0.055	0.01	0.15	1	Piecewise Linear (·, 0.05)
Disease-Free Survival, Surgery	P_s	0	0.24	0.25	0.55	1	Piecewise Linear (·, 0.05)
Fold-Increase in Disease-Free Survival	f_i	0	0	1.3	1.5	—	0.878 Gamma (6.392, 1)$^{0.216}$
Morbidity, Radiation	m_r	0	0.01	0.01	0.03	0.03	Uniform (0.01, 0.03)
Morbidity, Surgery	m_s	0	0.13	0.3	0.5	1	Piecewise Linear (·, 0.05)
Quality of Life Adjustment, Radiation	q_r	0	0.75	0.9	1	1	Piecewise Linear (·, 0.05)
Quality of Life Adjustment, Surgery	q_s	0	0.5	0.7	0.95	1	Piecewise Linear (·, 0.05)
Hospitalization, Surgery	T_s	0	0	5	24.5	—	4.645 Gamma (2.041, 1)$^{0.962}$

Min (Max) denote a parameter's minimum (maximum) feasible value; Lower (Upper) designate the lower (upper) bound of the parameter's plausible range; Base refers to the parameter's base value. We selected parameter distributions so that the mode corresponded to Base and 95% of the probability mass fell within [Lower, Upper]. A Piecewise Linear (·, 0.05) function is a density function linear between the parameter's breakpoints (i.e., its Min, Lower, Base, Upper, and Max values) with 95% of probability mass falling within [Lower, Upper] and 5% equally divided between [Min, Lower] and [Upper, Max]. The exceptions were the parameters m_r, which was uniformly distributed over its plausible range, and the parameters d_r, T_{sm} and T_{rm}, which were held fixed at base values for the analysis.

Table 2. EVPIs for parameter sets ξ

	d_{sr}	m_r	m_s	T_s	p_s	fi	q_s	d_s	q_r	p_r
p_r	2.3789	2.3789	2.3793	2.4811	1.9064	2.6579	2.9088	3.1880	3.5501	2.3789
q_r	1.2700	1.2724	1.2700	1.2708	1.6929	1.9879	1.9315	2.0246	1.2700	
d_s	0.7513	0.8079	0.8074	0.7805	1.4243	1.0614	1.4562	0.8081		
q_s	0.6140	0.6140	0.6149	0.7606	0.5402	0.8991	0.6140			
fi	0.2639	0.2652	0.2639	0.2686	0.5277	0.2639				
p_s	0.1627	0.1404	0.1406	0.1831	0.1404					
T_s	0.0028	0.0028	0.0028	0.0028						
m_s	0	0	0							
m_r	0	0								
d_{sr}	0									

EVPI(All) = 6.0660

Parameters the authors declared sensitive

Combinations for which authors performed SA

The table entry for row ξ_i and column ξ_j is EVPI(ξ_i, ξ_j); table entries for row ξ_i and column ξ_i depict EVPI(ξ_i). The optimal policy at base parameter values was Surgery Then Post-Operative Irradiation, which yielded an expected value of 108 additional quality adjusted weeks of survival.

CONCLUSION

The stability of the DM's preferred alternative to parametric variation has classically been addressed by SA measures based on threshold proximity, range of value, and probability of decision change. Only the latter approach adopts the Bayesian perspective that uncertainty should be probabilistically quantified. However, 'How far?', 'How much?' and 'How likely?' address only partial aspects of the sensitivity issue. The real issue is compound: 'How likely and with what effect?' To address both concerns, we recommend the use of information value as a sensitivity measure.

ACKNOWLEDGEMENTS

We would like to acknowledge the two anonymous referees for their comments. This work was supported by funding from the Decision, Risk and Management Science Program at the National Science Foundation.

REFERENCES

1. Briggs, A., Sculpher, M. and Buxton, M. Uncertainty in the economic evaluation of health care technologies: the role of sensitivity analysis. *Health Economics* 1994; **3**: 95–104.
2. Plante, D.A., Piccirillo, J.F. and Sofferman, R.A. Decision analysis of treatment options in pyriform sinus cancer. *Medical Decision Making* 1987; **7**: 74–83.
3. Dippel, D.W.J., Touw-Otten, F. and Habbema, J.D.F. Management of children with acute pharyngitis: a decision analysis. *Journal of Family Practice* 1992; **34**: 149–159.
4. Doubilet, P., Begg, C.B., Weinstein, M.C., Braun, P. and McNeil, B.J. Probabilistic sensitivity analysis using Monte Carlo simulation. *Medical Decision Making* 1985; **5**: 157–177.
5. Sonnenberg, F.A., Roberts, M.S., Tsevat, J., Wong, J.B., Barry, M. and Kent, D.L. Toward a peer review process for medical decision analysis models. *Medical Care* 1994; **32**: JS52–JS64.
6. O'Brien, B.J., Drummond, M.F., Labelle, R.J. and Willan, A. In search of power and significance: issues in the design and analysis of stochastic cost-effectiveness studies in health care. *Medical Care* 1994; **32**: 150–163.
7. Schneller, G.O. and Sphicas, G.P. On sensitivity analysis in decision theory. *Decision Sciences* 1985; **16**: 399–409.
8. Rios Insua, D. and French, S. A framework for sensitivity analysis in discrete multi-objective decision-making. *European Journal of Operational Research* 1991; **54**: 176–190.
9. Hughes, C. and Hughes, K.K. N-way sensitivity analysis: a verification and stability-assessment technique for completely subjective decision analysis trees. *Medical Decision Making* 1990; **10**: 68–74.
10. Merz, J.F., Small, M.J. and Fischbeck, P.S. Measuring decision sensitivity: a combined Monte

Carlo-logistic regression approach. *Medical Decision Making* 1992; **12**: 189–196.
11. Felli, J.C. and Hazen, G.B. Sensitivity analysis and the expected value of perfect information. *Medical Decision Making* 1998; **18**: 95–109.
12. Felli, J.C. and Hazen, G.B. Do sensitivity analyses really capture problem sensitivity? An empirical analysis based on information value. *Risk Decision and Policy* 1999; submitted, 1998.
13. Howard, R.A. Information value theory. *IEEE Transact Systems Science and Cybernetics* 1966; **SCC-2**: 22–26.
14. Howard, R.A. Value of information lotteries. *IEEE Transactions on Systems Science and Cybernetics* 1967; **SCC-3**: 24–60.
15. Wright, J.G. The minimal important difference: who's to say what is important? *Journal of Clinical Epidemiology* 1996; **49**:1221–1222.
16. Redelmeier, D.A., Guyatt, G.H. and Goldstein, R.S. Assessing the minimal important difference in symptoms: a comparison of two techniques. *Journal of Clinical Epidemiology* 1996; **49**: 1215–1219.
17. Redelmeier, D.A., Guyatt, G.H. and Goldstein, R.S. On the debate over methods for estimating the clinically important difference. *Journal of Clinical Epidemiology* 1996; **49**: 1223–1224.
18. Redelmeier, D.A. and Lorig, K. Assessing the clinical importance of symptomatic improvements. *Archives of Internal Medicine* 1993; **153**: 1337–1342.