Javelin Diagrams: A Graphical Tool for Probabilistic Sensitivity Analysis

James C. Felli

Naval Postgraduate School

Gordon B. Hazen

Northwestern University

Forthcoming in Decision Analysis.

Abstract

In order to demonstrate post-hoc robustness of decision problems to parameter estimates, analysts may conduct a *probabilistic sensitivity analysis*, assigning distributions to uncertain parameters and computing the probability of decision change. In contrast to classical threshold proximity methods of sensitivity analysis, no appealing graphical methods are available to present the results of a probabilistic sensitivity analysis. Here we introduce an analog of tornado diagrams for probabilistic sensitivity analysis, which we call *javelin diagrams*. Javelin diagrams are graphical augmentations of tornado diagrams displaying both the probability of decision change and the information value associated with individual parameters or parameter sets. We construct javelin diagrams for simple problems, discuss their properties, and illustrate their realistic application via a probabilistic sensitivity analysis of a seven-parameter decision analysis from the medical literature.

Introduction

Sensitivity analysis is today a crucial element in any practical decision analysis, and can play any of several roles in the decision analysis process. In the *basis development* phase of modeling (Howard 1983, 1988), sensitivity analysis can be used to guide model development: If decisions are insensitive to changes in some aspect of the model, then there is no need to model that particular aspect in more detail (Howard 1983, Watson and Buede 1987, Ch. 7). For instance, Clemen (1996, Ch.5) illustrates how *tornado diagrams* (e.g., Howard 1988) can be used to determine which deterministic variables have sufficient impact to be worth modeling probabilistically. Sensitivity analysis can also be used to give analysts *insight* into decision problems by identifying key variables (von Winterfeldt and Edwards 1986, Ch. 11; French 1986, pp. 342, 346). Sensitivity analysis may also be used to reduce the burden of assessing probabilities or utilities. For example, if a *strategy-region diagram* (e.g., Clemen 1996, Ch. 5) shows that the optimal decision remains so in a wide region of probabilities or utilities, then no detailed assessment is required. Examples are given by Clemen (1996, Ch. 5), Keeney and Raiffa (1976, pp. 100, 203), Watson and Buede (1987 p. 270) and von Winterfeldt and Edwards (1986, Ch. 11). Finally, in what Howard (1983) calls the *defensible stage*, sensitivity analysis may be used in a *post hoc* fashion (that is, after the analysis is complete) to demonstrate to supportive or skeptical audiences the *robustness* of the analysis, or to point out that a decision is a close call (von Winterfeldt and Edwards 1986, p. 401; French 1986, p. 252; Keeney and Raiffa 1976, p. 460).

Analysts have long recognized the dimensionality limitations of graphically based sensitivity analysis in portraying the robustness of a decision analysis to variations in underlying parameter estimates. If graphical methods allow at most 2- or 3-way sensitivity analyses, how can one be sure that a decision analysis is robust to the simultaneous variation of 20 to 30 parameters? Probabilistic sensitivity analysis was introduced to address this issue. In a probabilistic sensitivity analysis, the analyst assigns distributions to uncertain parameters and can thereby compute as a measure of robustness the probability of a change in the optimal alternative due to variation in arbitrarily many parameters.

In contrast to conventional methods of sensitivity analysis, which emphasize intuitively appealing graphical displays, the probabilistic sensitivity analysis approach is relatively devoid of graphical features. In this paper we introduce what might be termed the analog of tornado diagrams for probabilistic sensitivity analysis, which we call *javelin diagrams*. In a graphical display much like a tornado diagram, a javelin diagram displays not only the range of potential improvement offered by competing alternatives, but also the *probability of decision change* due to parametric variation and the *information value* associated with that variation. Although probabilistic analogues of tornado diagrams, javelin diagrams are intended to address post-hoc robustness issues, whereas tornado diagrams were designed for use in the basis formulation stage of a decision analysis.

Probabilistic sensitivity analysis was first adopted in medical decision analyses (Doubilet et al. 1985, Critchfield and Willard 1986, Critchfield, Willard and Connelly 1986). It has remained popular in medicine and health economics (Manning, Fryback and Weinstein 1996, Sisk et al. 1997, Goodman et al. 1999, Lord and Asante 1999, Ng et al. 1999, Pasta et al. 1999, Murakami and Ohashi 2001, Williams et al. 2001), and has seen application in civil and environmental engineering (Cawlfield and Wu 1993, Piggott and Cawlfield 1996, Lin et al. 1999). Because probabilistic sensitivity analysis requires that distributions be assigned to all parameters, there may be an additional assessment burden imposed; however, for conventional sensitivity analysis purposes, a responsible analysis will already have specified a best estimate and plausible range for each parameter. From there it is a mechanical procedure to fit an appropriately chosen distribution with, say, 95% of its probability mass in the plausible range and mean or mode equal to the best estimate, in which case no additional probability assessment is needed.

Conventional methods for sensitivity analysis, such as strategy-region diagrams or tornado diagrams, all rely implicitly on a *threshold proximity* view of sensitivity in which the analyst forms an intuitive judgment of sensitivity by examining the proximity of a parameter's base value to the nearest threshold of decision change. A similar viewpoint is prevalent in probabilistic sensitivity analysis, where the analyst computes the probability of a threshold crossing. In recent work (Felli and Hazen 1998, 1999), we have advocated what might be termed a *value-focused* view (Keeney 1992) for probabilistic sensitivity analysis, in which the analyst assesses sensitivity by computing the *value of information* for a parameter or parameter set. Empirical results indicate that in comparison with information value, threshold proximity approaches tend to significantly overestimate problem sensitivity (Felli and Hazen 1998, 1999).

All measures of post-hoc robustness attempt to quantify the degree to which parameter uncertainty may produce a change in the optimal choice. That is the point of view we adopt in this paper. Under this perspective, the sensitivity of the optimal *payoff* to parameter variation is unimportant unless it produces a change in the optimal *choice*. For example, if parameter variation cannot produce a change in optimal choice, then threshold proximity is nonexistent, the probability of decision change is zero, and the information value is also zero, *regardless* of how much the optimal payoff changes with parameter variation.

The paper is organized as follows. In section 1 we review tornado diagrams and introduce necessary terminology. In section 2, we move to probabilistic sensitivity analysis and introduce javelin diagrams. Section 3 addresses the general multi-parameter, multiple-alternative version of javelin diagrams and extends the discussion to incorporate non-neutral risk attitudes. In section 4, we apply the tool to a previously published medical decision analysis. We then close with a discussion of the relative advantages of javelin diagrams.

1. Tornado Diagrams

Consider the decision problem in figure 1, in which the probability p and the payoffs X, Y, Z are parameters, and the decision maker wishes to maximize expected payoff. Let $\Pi = (p, X, Y, Z)$ be the set of problem parameters, and let $\Pi_0 = (p_0, X_0, Y_0, Z_0)$ be the set of *base parameter values*, the estimated values of those parameters. We will denote the expected payoff of an alternative *a* as a function of its parameters Π by $E[V_a \mid \Pi]$. In this case, we have

$$E\left[V_{a_0} \middle| \Pi\right] = X$$

$$E\left[V_{a_1} \middle| \Pi\right] = pY + (1 - p)Z\ .$$

The alternative that maximizes expected payoff is a function $a_0(\Pi)$ of problem parameters. We let $a_0 = a_0(\Pi_0)$ be the optimal alternative given $\Pi = \Pi_0$ and will refer to a_0 as the *base optimal alternative* (BOA).

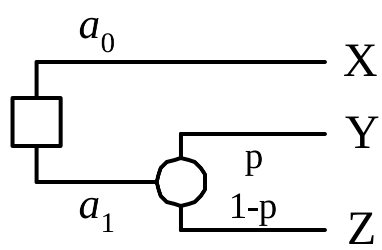

Figure 1. A simple decision problem with parameters p, X, Y and Z.

Base values and plausible ranges for the four parameters in this problem are shown in table 1. Using these data, we see that $\mathrm{E}\left[\mathrm{V}_{a_0}\middle|\Pi_0\right] = \$25$ and $\mathrm{E}\left[\mathrm{V}_{a_1}\middle|\Pi_0\right] = -\$20$, so a_0 is the BOA.

Parameter	Base	Minimum	Maximum
p	0.667	0	1
X	$25	-$200	$250
Y	-$133.33	-$400	$0
Z	$206.67	$0	$620

Table 1. Base parameter values and ranges for the decision problem in figure 1.

For the purposes of this illustration, we assume that successive cycles of model refinement are complete, and the analyst wishes to conduct a post-hoc robustness analysis on the parameters p, X, Y, and Z. A tornado diagram that accomplishes this is given in figure 2. The tornado diagram depicts the range of possible payoff gains

$$\mathrm{E}\left[\Delta\mathrm{V}_{a_1}\middle|\xi,\Pi_0\backslash\xi_0\right] = \mathrm{E}\left[\mathrm{V}_{a_1}\middle|\xi,\Pi_0\backslash\,\xi_0\right] - \mathrm{E}\left[\mathrm{V}_{a_0}\middle|\xi,\Pi_0\backslash\,\xi_0\right]$$

obtainable by varying the parameter ξ over its plausible range, for each of the four parameters $\xi \in \Pi$. Here, $\Delta\mathrm{V}_{a_1} = \mathrm{V}_{a_1} - \mathrm{V}_{a_0}$, and the notation $\Pi_0 \backslash \xi_0$ indicates the set Π_0 of base-value parameters with ξ_0 excluded. For example, when ξ = p, we have

$$\mathrm{E}\left[\Delta\mathrm{V}_{a_1}\middle|\mathrm{p},\Pi_0\backslash\,\mathrm{p}_0\right] = \mathrm{E}\left[\mathrm{V}_{a_1}\middle|\mathrm{p},\Pi_0\backslash\,\mathrm{p}_0\right] - \mathrm{E}\left[\mathrm{V}_{a_0}\middle|\mathrm{p},\Pi_0\backslash\,\mathrm{p}_0\right]$$

$= (pY_0 + (1 - p)Z_0) - X_0$

$= p(-133.33) + (1 - p)(206.67) - 25.$

$= 181.67 - 340\,p$

The range of this function of p is depicted as the p-bar in figure 2. The other three payoff gain functions are

$$E\left[\Delta V_{a_1} \middle| X, \Pi_0 \setminus X_0\right] = (p_0Y_0 + (1 - p_0)Z_0) - X = -20 - X$$

$$E\left[\Delta V_{a_1} \middle| Y, \Pi_0 \setminus Y_0\right] = (p_0Y + (1 - p_0)Z_0) - X_0 = 43.82 + 0.667Y$$

$$E\left[\Delta V_{a_1} \middle| Z, \Pi_0 \setminus Z_0\right] = (p_0Y_0 + (1 - p_0)Z) - X_0 = -113.93 + 0.333Z$$

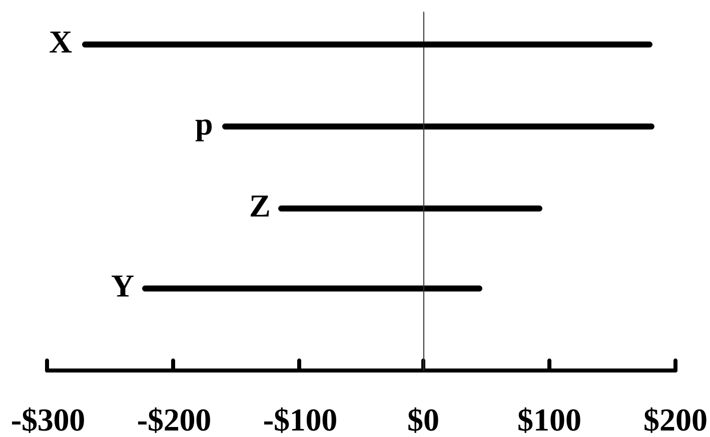

Figure 2. A tornado diagram in which deviations to the right of zero indicate potential gains over the payoff of the base optimal alternative a_0.

The quantity $E\left[\Delta V_{a_1} \middle| \xi, \Pi_0 \setminus \xi_0\right]$ is the payoff gain the decision maker could expect from a_1 over a_0 as a function of ξ. Using this formulation, zero becomes the natural reference point: the positive portion of a tornado bar for ξ corresponds to values of ξ for which the competing alternative is superior to the BOA. It is also possible to produce a tornado diagram that depicts the range of the alternative-a_1 payoffs $E\left[V_{a_1} \middle| \xi, \Pi_0 \setminus \xi_0\right]$ as a function of ξ, with reference point the alternative-a_0 payoff $E\left[V_{a_0} \middle| \Pi_0\right]$ (e.g., see the Eagle Airlines example in Chapter 5 of Clemen 1996). This format has the disadvantage that sensitivity to ξ is not properly depicted when $E\left[V_{a_0} \middle| \Pi\right]$ is also a function of ξ. For our example, either version of the tornado diagram would be adequate.

2. Javelin Diagrams

An alternate method of checking post-hoc robustness is to perform a probabilistic sensitivity analysis. It is easy to illustrate probabilistic sensitivity analysis using tornado diagrams. Suppose, for example, that the parameters $\Pi = (p, X, Y, Z)$ from our previous example are assigned beta densities as described in table 2. These densities are depicted graphically in figure 3. We assume no probabilistic dependence between the parameters.

Parameter	Base	Minimum	Maximum	Beta(α,β) densities α	β
p	0.667	0	1	9	4.5
X	\$25	-\$200	\$250	7	7
Y	-\$133.33	-\$400	\$0	3.15	1.575
Z	\$206.67	\$0	\$620	2.475	4.95

Table 2. Base values, ranges and distributions for the parameters of the decision problem in figure 1. All parameters are distributed as scaled beta densities with base values equal to expected values.

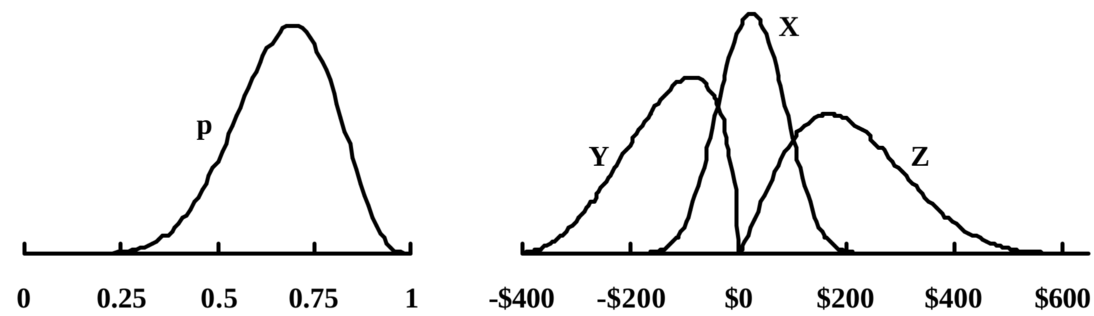

Figure 3. Parameter densities for the decision problem of figure 1.

Using these parameter densities, we can determine the distribution of each of the payoff gains $E\left[\Delta V_{a_1} \middle| \xi, \Pi_0 \setminus \xi\right]$ in the tornado diagram we created in figure 2. In simple cases, these distributions may be determined algebraically; in more complex cases, numerical methods such as Monte Carlo simulation may be required. We used Monte Carlo simulation to generate the payoff-gain densities for the decision problem of figure 1. These are presented in figure 4.

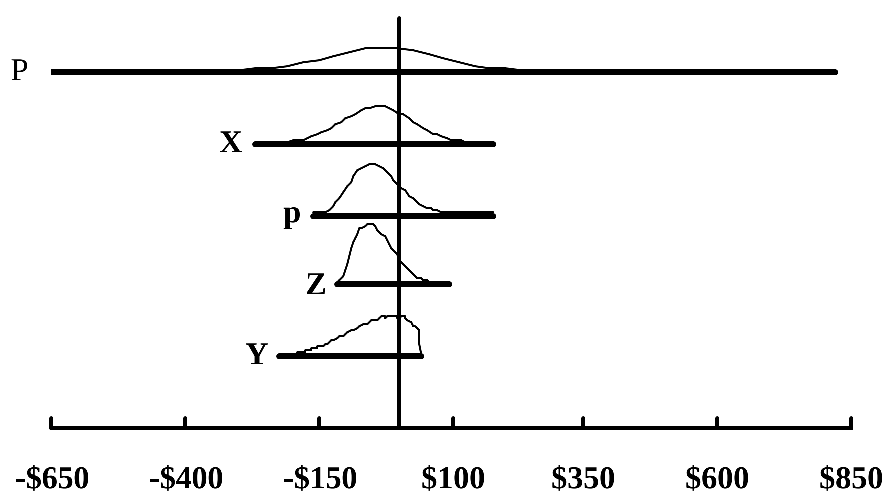

Figure 4. Payoff-gain densities for the decision problem of figure 1, imposed over the tornado bars of Figure 2. We also depict the payoff-gain density for the entire parameter set Π.

Another indicator of sensitivity in tornado diagrams is the absolute length of the tornado bar to the right of zero, which equals the *maximum payoff gain* due to variation in the corresponding parameter. For example, in figure 4, the maximum payoff gain due to variation in X is $179.89. The maximum payoff gain for Y is $43.82. (In line with our discussion in the introduction, it is important to notice we should not be interested in the *minimum* payoff gain because the BOA remains optimal in the region to the left of zero.)

A more revealing sensitivity indicator would be the *average positive payoff gain*, equal to

$$\mathrm{E}\left[\mathrm{E}\left[\Delta \mathrm{V}_{a_1}\middle|\xi,\Pi_0 \setminus \xi\right]^+\right],$$

and formed by using the parameter distribution to average the positive parts of the payoff gains. We do not average the payoff gain itself because parameter variation producing a negative payoff gain does not result in suboptimality for the base-optimal alternative, and therefore produces no decision change. We will use the more succinct term *expected improvement* to refer to average positive payoff gain. For many common decision problems (including the decision problem of figure 1), the expected improvement for a parameter ξ is equal (see Appendix) to the *expected value of perfect information on* ξ, which we denote EVPI_ξ.

As we have noted, in the decision problem of figure 1, we have

$$\mathrm{E}\left[\Delta \mathrm{V}_{a_1}\middle|\Pi\right] = \mathrm{pY} + (1-\mathrm{p})\mathrm{Z} - \mathrm{X}\,.$$

Therefore, the expected improvement for p, for example, is given by

$$\mathrm{E}\left[\mathrm{E}\left[\Delta \mathrm{V}_{a_1}\middle|\mathrm{p},\Pi_0 \setminus \mathrm{p}_0\right]^+\right] = \mathrm{E}\left[\left(\mathrm{pY}_0 + (1-\mathrm{p})\,\mathrm{Z}_0 - \mathrm{X}_0\right)^+\right].$$

Expected improvement is not available directly in the payoff-gain density diagrams of figure 4, although it may be calculated by multiplying the probability of decision change by the conditional expected payoff given a decision change (i.e., the center of gravity of the portion of the payoff-gain density to the right of zero).

For *javelin diagrams* as we now define them, both the expected improvement and likelihood of improvement may be viewed graphically. In a javelin diagram, instead of the density of each payoff gain $E\left[\Delta V_{a_1} \middle| \xi, \Pi_0 \setminus \xi\right]$, we graph the complementary cumulative distribution of this quantity; That is, we graph the function

$$G_\xi(\Delta v) = P(E\left[\Delta V_{a_1} \middle| \xi, \Pi_0 \setminus \xi\right] > \Delta v)$$

for $\Delta v \geq 0$. The javelin diagram consequent to figure 4 is shown in figure 5. The diagram gets its name from the graphical shaft formed by plotting the range of payoff gain and the head formed by the graph of $G_\xi(\Delta v)$ to the right of $\Delta v = 0$. The likelihood of improvement $G_\xi(0)$ is the height of the javelin head and the expected improvement is the area to the right of zero under the curve $G_\xi(\Delta v)$. In figure 5, the number to the left of the vertical line at zero for each ξ-bar is the probability $G_\xi(0)$ that the BOA loses optimality due to variation in ξ. The number to the right of the zero line is the expected improvement, equal to $EVPI_\xi$ in this case.

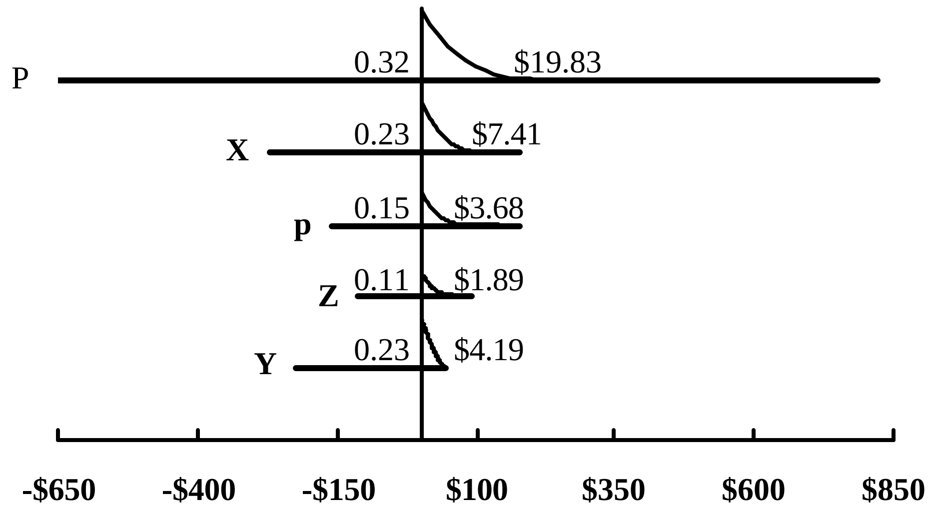

Figure 5. A javelin diagram for the decision problem of figure 1. The javelin curves indicate the probability that parametric variation yields payoff gain at least as great as that noted on the horizontal axis. Probabilities displayed to the left of each javelin head are the javelin height, equal to the probability of decision change. The values displayed to the right of each javelin are the areas under the javelin curves, equal to the expected value of perfect information of that parameter or parameter set.

The simultaneous display of expected improvement and likelihood of improvement is a key feature of javelin diagrams. The javelin diagram in figure 5 shows, for example, that: 1) there is a 32% chance that the competing alternative will outperform the BOA due to simultaneous variation in all parameters, and 2) by observing all parameters and adapting his choice as necessary, the decision maker could increase his payoff by an average of $19.83 over what he would expect from the BOA (i.e., $EVPI_{\Pi}$ = $19.83). We contend that expected improvement is a better sensitivity indicator than probability of decision change because it measures the *degree of impact* of a decision change rather than merely indicating its likelihood, and measures this impact in units the decision maker cares about – *payoff gains*. The javelin diagram provides intuitive support for this point of view. Compare the javelin for X in figure 5 with the javelin for Y: the larger javelin head for X gives the intuitive – and we argue, correct – impression that the problem is more sensitive to X than to Y, even though both parameters produce the same 23% likelihood of decision change.

How should one use the information provided in a javelin diagram? In this situation, the decision maker might reason as follows: In a problem in which the base optimal payoff is \$25, an expected improvement of \$19.83 = EVPI_Π seems substantial. Therefore, the problem seems jointly sensitive to all its parameters. Payoff gains of \$3 or greater seem significant, so the problem appears sensitive to the individual parameters p, X, Y as each has an EVPI value in excess of \$3. However, the problem may not be sensitive to the parameter Z individually, as its EVPI is only \$1.89. We state these assertions tentatively because they depend on a subjective declaration by the decision maker as to what constitutes a significant improvement in payoff.

3. Multiple Alternatives and Non-Neutral Risk Attitude

Multiple Alternatives

Both tornado diagrams and javelin diagrams may be readily adapted to the case of multiple alternatives. Suppose a_0 is the BOA and let $\Delta V_a = V_a - V_{a_0}$ be the difference between the payoff V_a under a and the payoff V_{a_0} under a_0. If ξ is a parameter or set of parameters, let $a_0(\xi)$ be the optimal alternative as a function of ξ when all other parameters $\Pi \setminus \xi$ lie at their base values. Here alternatives are chosen from a set A of feasible actions. Then the range of the payoff-gain function $\xi \to E\left[\Delta V_{a_0(\xi)} \middle| \xi, \Pi_0 \setminus \xi_0\right]$ may be graphed in a tornado diagram. Here the payoff gain $\Delta V_{a_0(\xi)} = V_{a_0(\xi)} - V_{a_0}$ is always nonnegative, so this tornado diagram would be truncated to the left at zero. To obtain an untruncated version let $a_0'(\xi)$ be the optimal alternative as a function of ξ

in the reduced feasible set $A \setminus a_0$. The payoff gain $E\left[\Delta V_{a_0'(\xi)} \middle| \xi, \Pi_0 \setminus \xi_0\right]$ may be positive or negative. The graph of the range of the function $\xi \rightarrow E\left[\Delta V_{a_0'(\xi)} \middle| \xi, \Pi_0 \setminus \xi_0\right]$ will yield an untruncated tornado diagram for payoff gains. Because $E\left[\Delta V_{a_0(\xi)} \middle| \xi, \Pi_0 \setminus \xi_0\right]$ is equal to $E\left[\Delta V_{a_0'(\xi)} \middle| \xi, \Pi_0 \setminus \xi_0\right]$ when the latter is nonnegative, the truncated and untruncated tornado diagrams will be identical to the right of zero.

Given a probability distribution for the parameter ξ, one can form a javelin diagram by first calculating the complementary cumulative distribution $G_\xi(\Delta v) = P\left(E\left[\Delta V_{a_0(\xi)} \middle| \xi, \Pi_0 \setminus \xi_0\right] > \Delta v\right)$ and then placing the graph of $G_\xi(\Delta v)$ for $\Delta v \geq 0$ over the corresponding untruncated tornado bar for the parameter ξ. The expected improvement $E_\xi\left[E\left[\Delta V_{a_0(\xi)} \middle| \xi, \Pi_0 \setminus \xi_0\right]\right]$ is the area to the right of zero under $G_\xi(\Delta v)$. The probability of decision change due to uncertainty in ξ is $G_\xi(0)$. As before, the expected improvement will often be equal to the information value of ξ (see Appendix).

Example: The non-equivalence of javelin and tornado diagrams

The javelin diagram is the natural extension of the tornado diagram for the purpose of incorporating prior probability distributions over parameters. Consequently, one might conjecture that when all parameters are uniformly distributed over their ranges, the javelin diagram conveys the same information and leads to the same sensitivity conclusions as the tornado diagram. The following example shows that this conjecture is false for three or more alternatives.

Consider the three-alternative example in figure 6. There are two parameters p, q with base values p_0, q_0 and ranges given in table 3. At the base values, the alternatives *a* = 0, 1, 2 have expected payoffs: $E\left[V_0 \mid p_0, q_0\right] = \$19.48$, $E\left[V_1 \mid p_0, q_0\right] = \$19.08$, $E\left[V_2 \mid p_0, q_0\right] = \$5.48$.

Therefore the BOA is $a_0 = 0$.

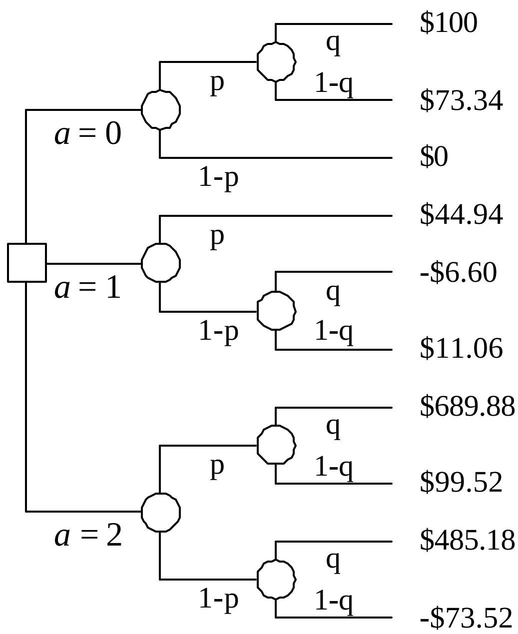

Figure 6. A hypothetical three-alternative decision problem with two probability parameters p and q.

Using the ranges and base values for p and q in table 3, we can construct the untruncated tornado diagram for payoff gains in figure 7a. If we assign distributions to p and q that are *uniform* over the ranges given in table 3, we can construct the javelin diagram in figure 7b.

Parameter	Base	Minimum	Maximum
p	0.26	0.01	0.51
q	0.06	0.005	0.115

Table 3. Parameter base values and ranges for the decision problem in figure 6.

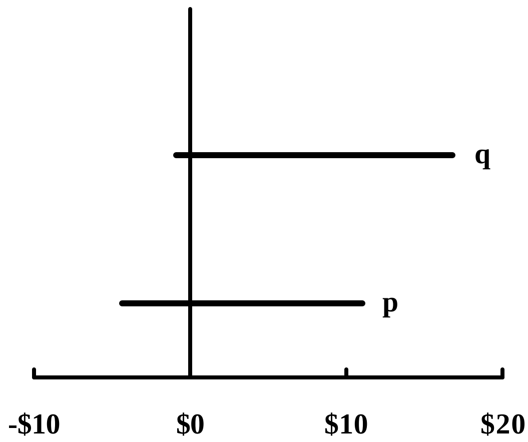

Figure 7a. Tornado diagram for the decision problem in figure 6. The bars represent the payoff gains over the BOA as p and q vary independently over their ranges.

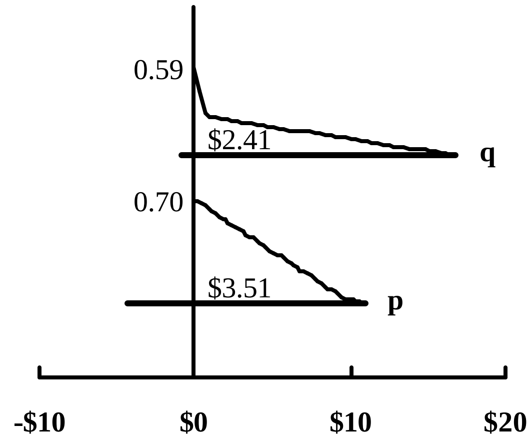

Figure 7b. Javelin diagram for the decision problem in figure 6. Numbers to the left denote the probability that a_0 loses optimality due to parametric variation; numbers to the right, the expected value of perfect parameter information for the parameter. The javelin curves indicate the probability that parametric variation yields payoff gain at least as great as on the horizontal axis.

The tornado diagram indicates that the problem is sensitive to both parameters, but more so to q than to p. The javelin diagram, on the other hand, indicates unambiguously that the problem is more sensitive to parameter p than to parameter q, exactly the reverse of the conclusion from the tornado diagram. Variation in p has a 70% chance of changing the optimal decision, with corresponding information value $\mathrm{EVPI}_p = \$3.51$. The problem is also sensitive to q but less so, with the probability of decision change equal to 54% and information value $\mathrm{EVPI}_q = \$2.40$. The range information in the tornado diagram does not adequately account for sensitivity in this case, even though both parameters were assigned uniform distributions. *Even with all parameters distributed uniformly, a javelin diagram need not convey the same sensitivity conclusions as a tornado diagram.*[1]

The reader may get some feeling for how this happens by examining figure 8, in which graphs of the functions $\xi \rightarrow \mathrm{E}\left[\Delta V_a \mid \xi, \Pi_0 \setminus \xi_0\right]$ appear for ξ = p, q and a = 0, 1, 2 in the decision problem of figure 6. The upper envelope in these graphs is the payoff-gain function $\xi \rightarrow \mathrm{E}\left[\Delta V_{a_0(\xi)} \middle| \xi, \Pi_0 \setminus \xi_0\right]$. The projection of each upper envelope onto the vertical axis (the vertical range) constitutes the nonnegative portion of the corresponding tornado bar. On the other hand, the average height of the upper envelope is $\mathrm{E}_\xi\left[\mathrm{E}\left[\Delta V_{a_0(\xi)} \middle| \xi, \Pi_0 \setminus \xi_0\right]\right]$ and is equal to the information value EVPI_ξ. This quantity is larger for ξ = p than for ξ = q, even though the vertical range is greater for ξ = q.

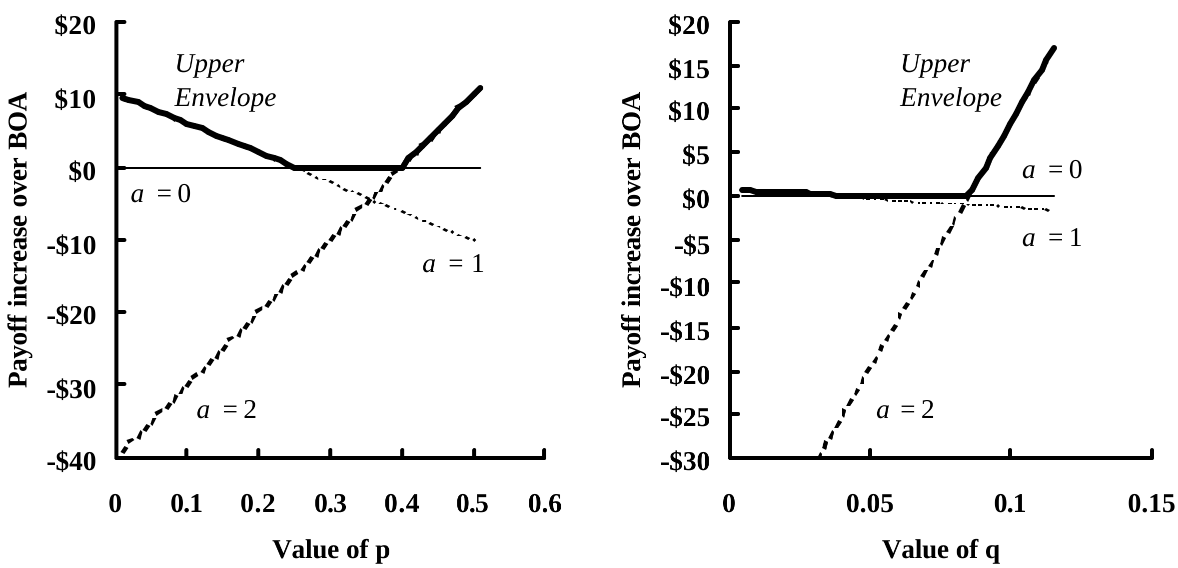

Figure 8. One-way sensitivity analysis graphs of the functions $\xi \rightarrow \mathrm{E}\left[\Delta V_a \mid \xi, \Pi_0 \setminus \xi_0\right]$ for ξ = p, ξ = q and a = 0, 1, 2 in the decision problem of figure 6. A tornado diagram depicts the vertical range of the upper envelope, which is larger for q than for p. However, because the parameters are distributed uniformly, information value is equal to the average height of the upper envelope and the probability of decision change is the proportion of the horizontal axis where the upper envelope exceeds zero. Both are greater for p than for q.

Non-Neutral Risk Attitude

We have introduced the javelin diagram as a sensitivity analysis tool for the cases in which payoff is monetary and the decision maker is risk-neutral. In fact no such restrictions are necessary: The notion of information value as expected improvement and the corresponding javelin display can be extended to the case of non-neutral risk attitude and even to the case in which there is no underlying monetary attribute. Because the approach we take is unorthodox – perhaps even heterodox to some – we present it here in detail.

Suppose a utility function u(x) has been assessed over problem outcomes x, and let U_a be a random variable equal to the utility achieved under alternative *a*. Let *a** be the alternative maximizing expected utility $E[U_a]$, and let $a^*(\xi)$ be the alternative *a* maximizing the conditional expected utility $E[U_a \mid \xi]$ given the value of parameter ξ. Define

$$EU_\xi = E_\xi[E[\, U_{a^*(\xi)} \mid \xi]],$$

the expected utility of learning the value of ξ. Similarly, let $EU_\emptyset = E[\, U_{a^*}\,]$ be the optimal expected utility under no further information.

Larger values of EU_ξ indicate greater problem sensitivity to ξ, so the quantities EU_ξ may be used to rank problem sensitivity to different parameter sets ξ. This is the first role of any sensitivity measure. The analyst may find it helpful to form the *expected improvement in utility*,

$$EUI_\xi = EU_\xi - EU_\emptyset.$$

However, this measure is merely for convenience. Although it is possible (see below) and even desirable to assign an information theoretic meaning to EUI_ξ, there is no need to do so for the purposes of sensitivity analysis.

We require a second role of any sensitivity measure, namely the declaration (yes/no) of problem sensitivity to a set ξ of parameters. To this end, let $\tilde{x}^*$ be the random outcome corresponding to the optimal alternative *a**. We require the decision maker or analyst to declare *an outcome having minimum significant improvement*, that is, a (possibly random) outcome $\tilde{x}^+$ having the properties that: 1) $\tilde{x}^+$ is preferred to $\tilde{x}^*$, 2) $\tilde{x}^+$ is *significantly* preferred to $\tilde{x}^*$, and 3) no other random or nonrandom outcome $\tilde{x}$ that is less preferred than $\tilde{x}^+$ can be significantly preferred to $\tilde{x}^*$. Declaring an outcome $\tilde{x}^+$ having minimum significant improvement over $\tilde{x}^*$ is a completely subjective judgment by the analyst or decision maker, and is not a property of the assessed utility function u. Nevertheless, the outcome $\tilde{x}^+$ having minimum significant improvement over $\tilde{x}^*$ induces a *minimum significant increase* Δu^+ *in utility,* given by

$$\Delta u^+ = E[u(\tilde{x}^+)] - E[u(\tilde{x}^*)].$$

The value Δu^+ of the minimum significant increase in utility is intrinsic to the particular decision problem at hand, as it depends on how the utility function is scaled and the optimal random outcome $\tilde{x}^*$. It is not possible to declare a minimum significant improvement in utility valid across all decision problems. Nevertheless, for the problem at hand, no outcome $\tilde{x}$ whose improvement in utility is less than Δu^+ will constitute a significant improvement over $\tilde{x}^*$.

We propose to extend the notion of significant improvement to information sources, and say that perfect information about a parameter set ξ *constitutes a significant improvement* if the expected utility EU_ξ of learning ξ exceeds the expected utility $E[u(\tilde{x}^+)]$ of $\tilde{x}^+$, or equivalently, if the expected improvement in utility EUI_ξ by learning ξ exceeds the minimum significant utility improvement Δu^+. We propose to declare a decision problem sensitive to a parameter set ξ if learning the value of ξ constitutes a significant improvement.

Although this approach to probabilistic sensitivity analysis has an information-theoretic basis, there is no requirement that the expected improvement in utility EUI_ξ itself have an information theoretic meaning. Nevertheless, EUI_ξ can in fact be interpreted as a form of information value, as we shall now explain. Suppose we let x^- be an outcome (call it the *negative outcome*) having utility $u(x^-)$ strictly smaller than optimal expected utility $EU_\varnothing$. Define the *probability price* PP_ξ of the parameter set ξ to be the largest probability of the negative outcome x^- one is willing to accept in order to obtain perfect information about ξ. The probability price of ξ is a measure of information value, and satisfies

$$PP_\xi u(x^-) + (1 - PP_\xi)EU_\xi = EU_\varnothing.$$

Rearranging and using $EUI_\xi = EU_\xi - EU_\varnothing$, we have

$$\frac{PP_\xi}{1 - PP_\xi} = \frac{EUI_\xi}{EU_\varnothing - u(x^-)},$$

We see therefore that the expected improvement in utility EUI_ξ is proportional to the odds price $\frac{PP_\xi}{1 - PP_\xi}$, that is, EUI_ξ is proportional to the largest odds of the negative outcome x^- one would accept to learn ξ. Rescaling utility if necessary, we can choose x^- so that $EU_\varnothing - u(x^-) = 1$, in which case EUI_ξ is the largest odds of a unit reduction of utility from $EU_\varnothing$ that one would accept to learn ξ.

Probability price and its relationship to expected improvement in utility were discussed by Hazen and Sounderpandian (1999). When outcomes are monetary and risk attitude is not neutral, the conventional definition is that information value $EVPI_\xi$ is the most one would be willing to pay to learn ξ, that is, $EVPI_\xi$ satisfies

$$E[u(x^* - EVPI_\xi) \mid \xi] = EU_\varnothing .$$

Hazen and Sounderpandian show that when risk attitude is constant (i.e., a linear or exponential utility function), then expected improvement in utility EUI_ξ and conventionally defined information value $EVPI_\xi$ rank information sources identically, but otherwise, they may differ. The advantage to using EUI_ξ is that it is computationally and analytically simpler when risk attitude is not constant, and it generalizes to situations in which outcomes are not monetary.

In the preceding discussion we have glossed over a practical detail involving base estimates for parameters. The optimal expected utility may be written

$$EU_\varnothing = E[U_{a*}] = E_\Pi[E[U_{a*} \mid \Pi]].$$

In the common special case in which $E[U_{a*} \mid \Pi]$ is a *multilinear* function of the parameters Π (i.e., linear in each component of Π) and parameters are independent, then we have

$$EU_\varnothing = E[U_{a*} \mid \bar{\Pi}]$$

where $\bar{\Pi} = E[\Pi]$ is the mean of Π. If the analyst has taken the base optimal level Π_0 of parameters Π to be the mean $\bar{\Pi}$, then it follows that a^* is the base optimal alternative a_0 obtained by maximizing $E[U_a \mid \Pi_0]$. Similarly, for any parameter set ξ, $a^*(\xi)$ will equal $a_0(\xi)$. (See the Appendix for details.) However, if these conditions fail, that is, if parameters are not independent, if multilinearity does not hold, or if base parameter levels are not equal to their mean values, then a^* may differ from a_0, and $a^*(\xi)$ may differ from $a_0(\xi)$. When this occurs, and the analyst computes expected improvements based on a_0 and $a_0(\xi)$ instead of a^* and $a^*(\xi)$, then the information-theoretic basis for the procedure is only approximately maintained.

4. A Realistic Example

As an example of constructing and interpreting a javelin diagram, we will revisit the problem of management of suspected giant cell arteritis posed by Buchbinder and Detsky (1992), hereafter referred to as B&D. This problem illustrates the use of javelin diagrams on a realistic-size problem, and the use of expected improvement in utility as a sensitivity measure, as discussed in the previous section.

B&D considered four alternatives in this problem: (A) Treat None, (B) Biopsy & Treat Positive, (C) Biopsy & Treat All, (D) Treat All. The structure we provide below is the same as B&D with the exception of the notation we adopted for ease of exposition. The four alternatives are presented in figures 9a-d.

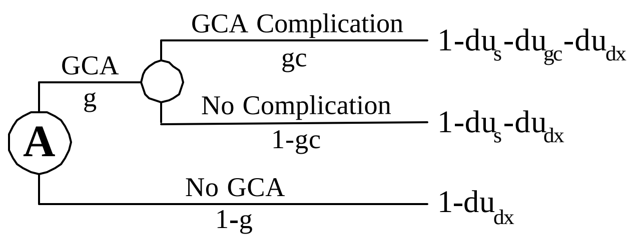

Figure 9a. The decision tree for the Treat None alternative.

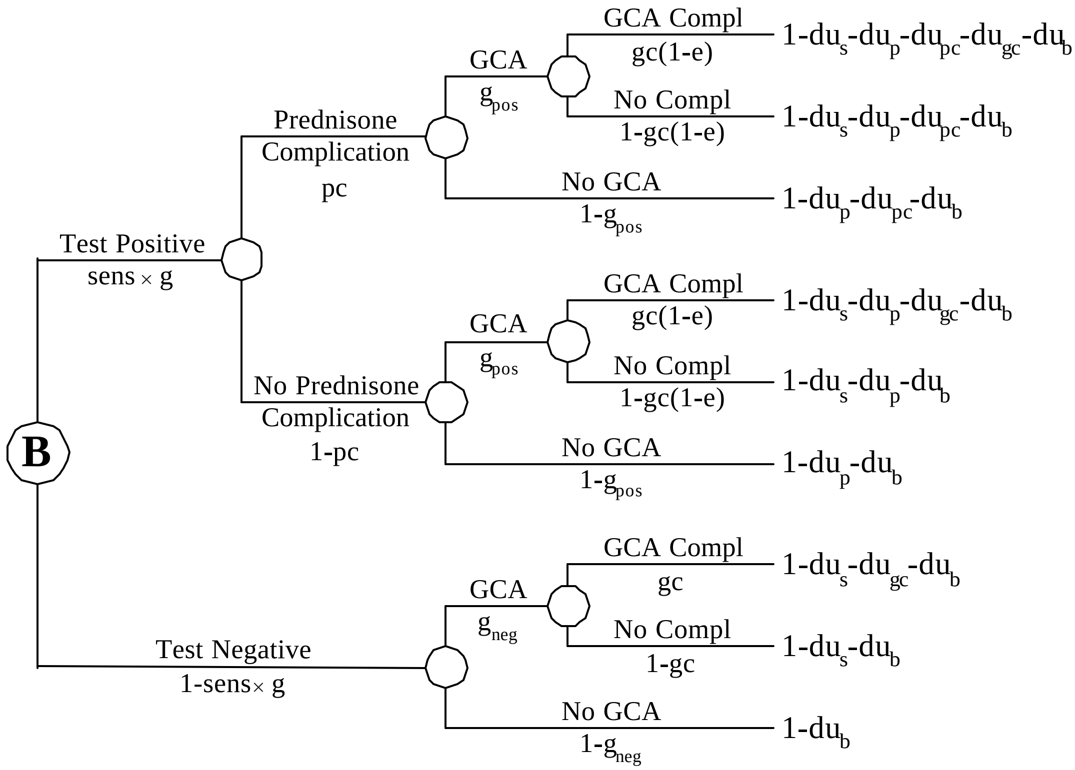

Figure 9b. The decision tree for the Biopsy & Treat Positive alternative. The probability of a positive test is equal to $\text{sens} \times \text{g}$. The terms g_{pos} and g_{neg} refer to the probability of having GCA given a positive or negative test result. Because the specificity of the test is 1, $g_{pos} = 1$ and $g_{neg} = (1 - \text{sens})\,g / (1 - \text{sens} \times g)$.

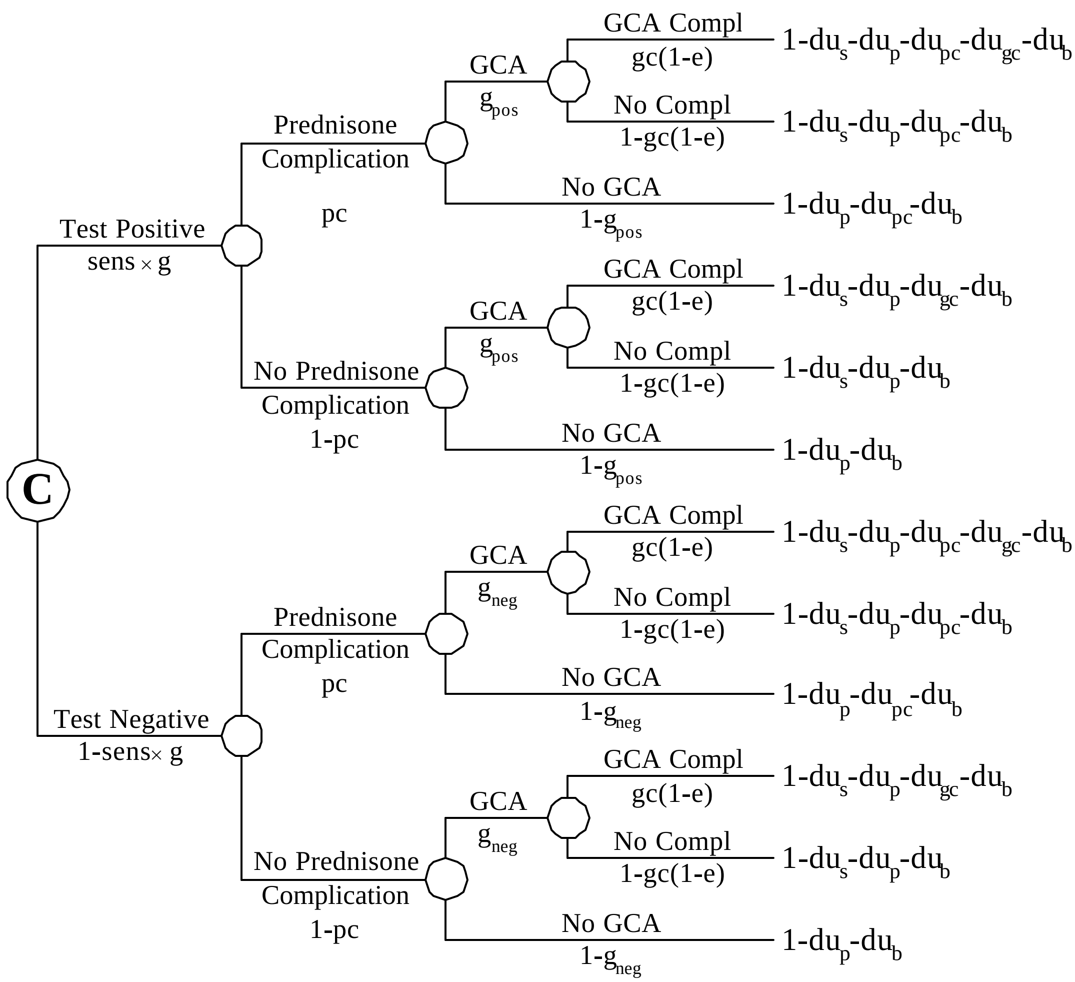

Figure 9c. The decision tree for the Biopsy & Treat All alternative. The probability of a positive test is equal to $\text{sens} \times \text{g}$. The terms g_{pos} and g_{neg} refer to the probability of having GCA given a positive or negative test result. Because the specificity of the test is 1, $g_{pos} = 1$ and $g_{neg} = (1 - \text{sens})\,\text{g} / (1 - \text{sens} \times \text{g})$.

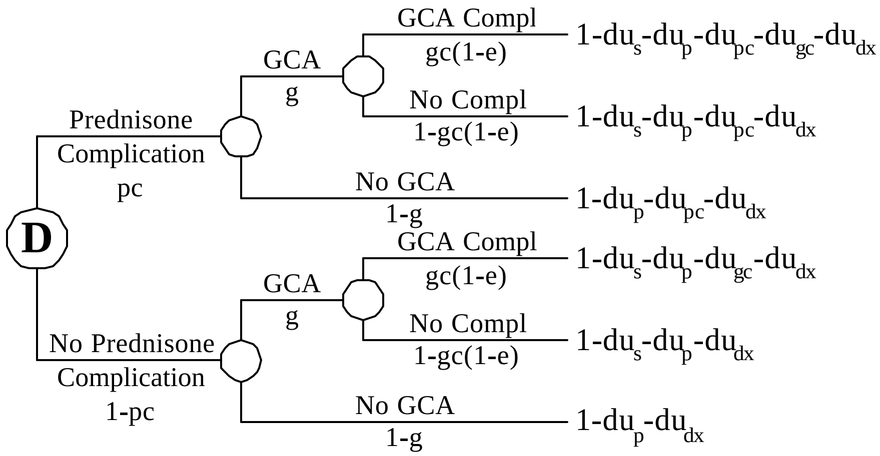

Figure 9d. The decision tree for the Treat All alternative.

Table 4 provides a summary of the parameter values used by B&D in their analysis and the range of values they used for a sensitivity analysis with the probability of having giant cell arteritis (GCA) set at 0.8. The reader may note that the expected utilities $E[U_a \mid \Pi]$ are all multilinear in the parameters Π. With all parameters held at base value, Biopsy & Treat Positive was optimal with an expected payoff 0.837 on a utility scale from -0.325 (all disutilities present) to 1 (perfect health). The expected utilities of the four treatment alternatives at base parameter values are provided in table 5. Note that all utilities in this analysis were physician assessed. The analysis therefore takes they perspective of the "benevolent physician".

The threshold column in table 4 notes the value of the parameter required to cause the BOA (Biopsy & Treat Positive) to lose optimality when all other parameters remain fixed at their base values. Except for the beta density parameters, B&D provided the data in table 4. We assumed parameters to be probabilistically independent and determined the beta density parameters by fitting beta distributions to B&D data, using the base values as mean values and capturing 95% of the probability mass in the range defined by the parameter's minimum and maximum plausible values. We let the remaining 5% spill equally to either side of this range if possible (that is, if the upper plausible limit was less than one and the lower limit greater than zero – otherwise we filled the plausible range with 97.5 or 100% probability mass, as appropriate).

Parameters	Symbol	Base	Minimum	Maximum	Threshold	Beta (a ,b) a	b
P[having GCA]	g	0.8	-	-	-	-	-
P[developing severe complications of GCA]	gc	0.3	0.05	0.5	0.31	4.719	11.011
P[developing severe iatrogenic side effects]	pc	0.2	0.05	0.5	0.19	2.647	10.589
Efficacy of high dose prednisone	e	0.9	0.8	1	0.94	27.787	3.087
Sensitivity of temporal arterty biopsy	sens	0.83	0.6	1	0.82	7.554	1.547
Specificity of temporal arterty biopsy	spec	1	-	-	-	-	-
D(major complication from GCA)	du_{gc}	0.8	0.3	0.9	0.84	27.454	6.864
D(Prednisone therapy)	du_p	0.08	0.03	0.2	0.08	4.555	52.380
D(major iatrogenic side effect)	du_{pc}	0.3	0.2	0.9	0.28	15.291	35.680
D(having symptoms of GCA)	du_s	0.12	-	-	-	-	-
D(having a temporal artery biopsy)	du_b	0.005	-	-	-	-	-
D(not knowing the true diagnosis)	du_{dx}	0.025	-	-	-	-	-

Table 4. The data used by B&D in their analysis. The minimum and maximum values depict each parameter's range for sensitivity analysis. The threshold value indicates the parameter value at which the BOA lost optimality to another alternative. The values $P[\cdot]$ and $D(\cdot)$ designate the probability of an event and the disutility associated with an event. B&D assigned $du_S = 0$ when treatment was given because the effectiveness of therapy completely alleviated the patient's symptoms.

Treatment Alternative	EU at Base Values
Treat None	0.687
Biopsy & Treat Positive	0.837
Biopsy & Treat All	0.836
Treat All	0.816

Table 5. The expected utilities of the four alternative treatments for suspected giant cell arteritis when all parameters are set to their base values in table 4. With all parameters at their base values, the B&D utility scale ranged from -0.325 (all disutilities present) to 1 (perfect health).

The close proximity of threshold values to parameter base values implies that a small parameter value deviation away from base value could cause the BOA to lose optimality. Based on this, B&D, in their post-hoc robustness analysis, declared the BOA sensitive to the seven non-constant parameters in table 4 when g = 0.8.

To check these sensitivity conclusions, we constructed a tornado diagram and javelin diagram. Our tornado diagram depicting utility improvements is provided in figure 10. The optimality of the BOA not only appears to be clearly sensitive to variation in the *sens* parameter, but extremely susceptible to variation in du_p. The javelin diagram corresponding to the beta densities in table 4 is provided in figure 11a. The javelin heads were determined via Monte Carlo simulation. Figure 11b is a rescaled version of figure 11a that gives a better view of the javelin heads.

Areas under the javelin heads are expected improvement *in utility*, discussed in the previous section. For this analysis, an expected improvement in utility of Δu for parameter ξ indicates that the possibility of learning ξ is worth a $\frac{\Delta u}{1+\Delta u}$ chance of falling one unit in utility from the base optimal level 0.837. A drop of one unit in utility from 0.837, that is, from 0.837 to −0.163, amounts to a total utility decrement of 1 − (−0.163) = 1.163 from the well state, approximately equivalent here to the simultaneous occurrence of major complications from GCA and major iatrogenic side effects.

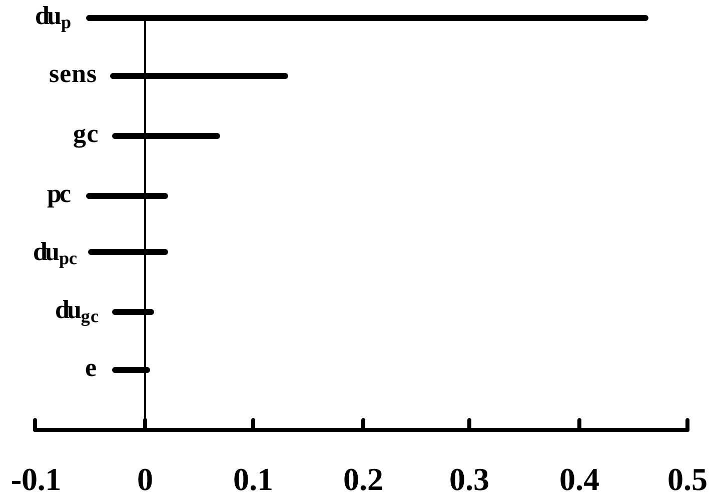

Figure 10. The payoff-gain tornado diagram for the giant cell arteritis problem.

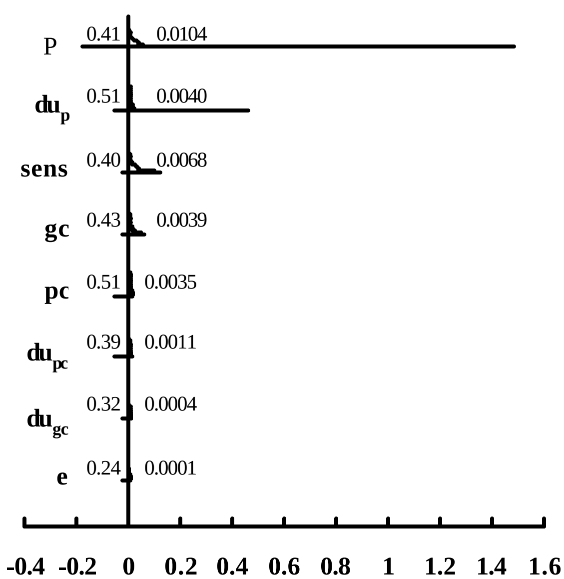

Figure 11a. A javelin diagram for the giant cell arteritis problem.

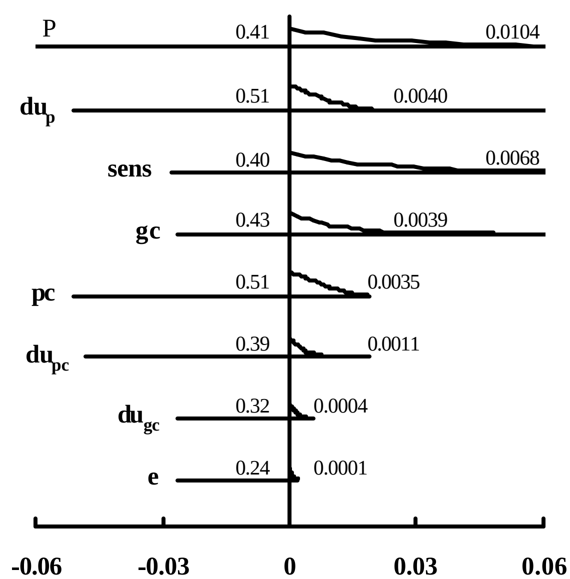

Figure 11b. Figure 11a enlarged to better show the javelin heads.

Figures 11a and b substantiate the results of the B&D threshold analyses in that the probability of another alternative yielding a higher expected utility than the BOA is high. However, the all-parameter information value is a utility improvement of merely 0.0104, indicating that learning the true values of all parameters is worth taking at most a $\frac{0.0104}{1+0.0104} = 1.03\%$ chance at the simultaneous occurrence of major complications from GCA and major iatrogenic side effects. The significance of a utility improvement of 0.0104 over the base-optimal utility 0.837 is a subjective judgment. However, it is worth noting that the smallest utility decrement noted by B&D in their analysis was the disutility for having a temporal artery biopsy, equal to 0.005. If one could eliminate the disutility of temporal artery biopsy from the optimal policy *biopsy and treat positive*, its expected utility would increase from 0.837 to 0.842. Presumably this constitutes a significant increase, as otherwise it is unlikely B&D would have included this disutility in their model. Moreover, because this was included and no smaller values were considered, we may be justified in assuming that 0.005 is the smallest utility increment B&D considered significant. By this light, the utility improvement 0.0104 for Π would be considered significant, as would the utility increment for *sens*, but no other individual parameter would possess significant information value. Note in particular that while variation in *gc* was likely to undermine the optimality of the BOA 43% of the time, its information value was only 0.0039. Where the tornado diagram of figure 10 illustrates that the BOA is sensitive to variation in *gc*, the small information value suggests otherwise. Moreover, although the BOA appears vastly more sensitive to du_p than *sens* based on the potential payoff gains mapped out in figure 10, the opposite is actually the case as $\mathrm{EUI}_{sens} > \mathrm{EUI}_{du_p}$ (0.0068 versus 0.0040), and the problem does not appear sensitive to du_p in information value terms. This is typical of what we have found in our reexamination of many published sensitivity analyses (Felli and Hazen 1998, 1999), namely that threshold proximity and probability of decision change frequently overestimate problem sensitivity compared to information value.

As in any probabilistic sensitivity analysis, the sensitivity conclusions obtained depend to some degree on the selected parameter distributions. However, in our experience, the shape of the distribution has only a small effect on probability or information-value conclusions. For example, in our re-analysis of 25 published decision analyses (Felli and Hazen 1999), there was little qualitative difference in probability or information-value conclusions when we replaced all parameter distributions by uniform distributions over the plausible parameter range. The changes in expected improvement consequent to using uniform rather than beta densities in the B&D problem are provided in table 6.

	Parameters							
Density Type	**e**	**sens**	**gc**	**pc**	**du_{gc}**	**du_p**	**du_{pc}**	P
Beta	0.0001	**0.0068**	0.0039	0.0035	0.0004	0.0040	0.0011	**0.0104**
Uniform	0.0003	**0.0096**	0.0038	0.0021	0.0001	0.0021	0.0003	**0.0267**

Table 6. Information values are noted for parameters in the B&D problem using two different distribution assumptions. The beta values were obtained using the beta(α,β) densities noted in table 4. The uniform values resulted from using uniform(A,B) densities, where A and B were the minimum and maximum parameter values listed in table 4. Bold-faced values exceed 0.005.

Although there is deviation in information values depending on the choice of parameter density, the optimality of the BOA in the B&D problem appears sensitive only to variation in *sens* and Π (based on the assumed demarcation point of 0.005 discussed earlier). The sensitivity conclusions the problem remain the same regardless of whether parameter values are distributed according the beta densities noted in table 4 or uniformly distributed over their plausible ranges.

5. Summary

We have introduced javelin diagrams as a graphical aid for displaying results from a probabilistic sensitivity analysis. Like tornado diagrams, javelin diagrams provide the analyst with a compact display of payoff range information as a parameter or set of parameters are allowed to vary. In addition, javelin diagrams also provide the analyst with probabilistic and information value-based indicators of decision sensitivity.

For ease of exposition, we assumed parametric independence in our examples. As it happens, correlations between parameters are rarely accounted for in conventional sensitivity analyses, even when they are well understood. For example, sensitivity and specificity of diagnostic tests are often treated as independent parameters for sensitivity analysis purposes despite their correlation through the thresholds used to declare positive test outcomes (Littenberg and Moses 1993). Some analysts build correlation directly into the model structure by specifying functional relationships between parameters. The Buchbinder and Detsky example we cite in section 4 illustrates one such method common in the medical literature: employment of an efficacy parameter when drug and/or treatment choice can affect the likelihood of specific, downstream events. Dependencies built into model structure are accounted for by all conventional sensitivity analysis tools, as well as javelin diagrams. This may well be the best way to account for probabilistic dependencies, as more explicit information on the joint densities of parameters is rarely available.

As is well known (e.g., Howard 1983), value-of-information computations can serve to guide the analyst's model refinement and information acquisition choices in the decision-analysis modeling cycle. There is no reason why a javelin diagram could not be used at these earlier stages as well as the post-hoc robustness stage we discuss in this paper. A small qualitative difference would occur for discrete variables, for which the javelin heads appear as step functions, but otherwise the same formulas and interpretations apply.

References

Buchbinder, R. and Detsky, A.S. 1992. Management of suspected giant cell arteritis: a decision analysis. Journal of Rheumatology 19 (8), 1220-1228.

Cawlfield, J.D. and Wu, M.C. 1993. Probabilistic sensitivity analysis for one-dimensional reactive transport in porous-media. Water Resources Research, 29 (3), 661-672.

RT Clemen, Making Hard Decisions: An Introduction to Decision Analysis, Duxbury Press, 1996.

Critchfield, G.G. and Willard, K.E. 1986. Probabilistic analysis of decision trees using Monte Carlo simulation. Medical Decision Making, 6 (2), 85-92.

Critchfield, G.G., Willard, K.E., Connelly, D.P. 1986. Probabilistic analysis methods for general decision models. Computers and Biomedical Research, 19, 254-265.

Doubilet, P., Begg, C.B., Weinstein, M.C., Braun, P., McNeil, B.J. 1985. Probabilistic sensitivity analysis using Monte Carlo simulation. Medical Decision Making, 5 (2), 157-177.

Felli, J.C. and Hazen, G.B. 1998. Sensitivity analysis and the expected value of perfect information. Medical Decision Making, 18 (1), 95-109.

Felli, J.C. and Hazen, G.B. 1999. Do sensitivity analyses really capture problem sensitivity? An empirical analysis based on information value. Risk, Decision and Policy, 4 (2), 79-98.

Goodman, C.A., Coleman, P.G., Mills, A.J. 1999. Cost-effectiveness of malaria control in sub-Saharan Africa. Lancet, 354 (9176), 378-385.

Hamed, M.M. 1999. Probabilistic sensitivity analysis of public health risk assessment from contaminated soil. Journal of Soil Contamination, 8 (3), 285-306.

Hazen, G.B. and J. Sounderpandian. 1999. Lottery acquisition versus information acquisition: prices and preference reversals. Journal of Risk and Uncertainty 18 (2), 125-136.

Howard, R.A. 1983. The evolution of decision analysis. In R.A. Howard and J.E. Matheson 1989: The Priciples and Applications of Decision Analysis. Strategic Decisions Group, Menlo Park, California.

Howard, R.A. 1988. Decision analysis: practice and promise. Management Science 14 (6), 679-695.

Keeney, R.L. 1992. *Value-focused thinking*. Harvard University Press, Cambridge, MA.

RL Keeney and H Raiffa, *Decisions with Multiple Objectives: Preferences and Value Trade-offs*, John Wiley and Sons 1976, and Cambridge University Press, 1993.

Lin, P.S., Chang, C.W., Juang, C.H. 1999. Probabilistic sensitivity analysis for one-dimensional solute transport. Civil Engineering and Environmental Systems, 17 (1), 39-61.

Littenberg, B. and Moses, L.E. 1993. Estimating diagnostic accuracy from multiple conflicting reports: a new meta-analytic method. Medical Decision Making, 13 (4), 313-321.

Lord, J. and Asante, M.A. 1999. Estimating uncertainty ranges for costs by the bootstrap procedure combined with probabilistic sensitivity analysis. Health Economics, 8 (4), 323-333.

Manning, W.G., Fryback, D.G., and Weinstein, M.C. 1996. Reflecting uncertainty in cost-effectiveness analysis. In *Cost-Effectiveness in Health and Medicine*, edited by M.R. Gold, J.E. Siegel, L.B. Russell, and M.C. Weinstein, Oxford University Press.

Murakami, Y. and Ohashi, Y. 2001. Projected number of diabetic renal disease patients among insulin-dependent diabetes mellitus children in Japan using a Markov model with probabilistic sensitivity analysis. International Journal of Epidemiology, 30 (5), 1078-1083.

Ng, A.K., Weeks, J.C., Mauch, P.M., Kuntz, K.M. 1999. Decision analysis on alternative treatment strategies for favorable-prognosis, early-stage Hodgkin's disease. Journal of Clinical Oncology, 17 (11), 3577-3585.

Piggott, J.H. and Cawlfield, J.D. 1996. Probabilistic sensitivity analysis for one-dimensional contaminant transport in the vadose zone. Journal of Contaminant Hydrology, 24 (2), 97-115.

Pasta, D.J., Taylor, J.L., Henning, J.M. 1999. Probabilistic sensitivity analysis incorporating the bootstrap: An example comparing treatments for the eradication of Helicobacter pylori. Medical Decision Making, 19 (3), 353-363.

Sisk, J.E., Moskowitz, A.J., Whang, W., Lin, J.D., Fedson, D.S., McBean, A.M., Plouffe, J.F., Cetron, M.S., Butler, J.C. 1997. Cost-effectiveness of vaccination against pneumococcal bacteremia among elderly people. Journal of the American Medical Association, 278 (16), 1333-1339.

von Winterfeldt, D. and Edwards, W. 1986. Decision Analysis and Behavioral Research. Cambridge University Press.

Watson, S.R. and Buede, D.M. 1987. Decision Synthesis: The Principles and Practice of Decision Analysis. Cambridge University Press.

Williams, P., Dowson, A.J., Rapoport, A.M., Sawyer, J. 2001. The cost effectiveness of stratified care in the management of migraine. Pharmacoeconomics, 19 (8), 819-829.

Appendix: Computing for javelin diagrams and information value

Let $a_0(\xi)$ be the optimal alternative as a function of ξ when all other parameters $\Pi \setminus \xi$ are held fixed at their base values $\Pi_0 \setminus \xi_0$. Computation of the expected improvement quantity

$$\mathrm{EVPI}_{\xi}^{0} = \mathrm{E}_{\xi}\left[\mathrm{E}\left[\Delta \mathrm{V}_{a_0(\xi)} \mid \xi, \Pi_0 \setminus \xi_0\right]\right]$$

and the javelin curve

$$\mathrm{G}_{\xi}(\Delta \mathrm{v}) = \mathrm{P}\left(\mathrm{E}\left[\Delta \mathrm{V}_{a_0(\xi)} \mid \xi, \Pi_0 \setminus \xi_0\right] > \Delta \mathrm{v}\right) \qquad \Delta \mathrm{v} \geq 0$$

is a straightforward exercise in Monte Carlo simulation, since the quantity $\mathrm{E}\left[\Delta \mathrm{V}_{a_0(\xi)} \mid \xi, \Pi_0 \setminus \xi_0\right]$ as a function of the random variable ξ is readily available from the decision problem formulation. The simulation can handle arbitrary parameter subsets $\xi \subset \Pi$, and if desired, Monte Carlo approximations for Π and multiple subsets ξ of Π can be calculated in a single simulation run.

As is well known, the mean of the non-negative random variable $\mathrm{E}\left[\Delta \mathrm{V}_{a_0(\xi)} \mid \xi, \Pi_0 \setminus \xi_0\right]$ is the area under its complementary cumulative distribution $\mathrm{G}_{\xi}(\Delta \mathrm{v})$. Therefore, the area under the javelin curve is EVPI_{ξ}^{0}.

If the regions $\{\xi \mid a_0(\xi) = a\}$ for each alternative a can be expressed as simple inequalities in ξ (as is often possible for single parameters ξ), then it may be possible to obtain closed-form expressions for EVPI_{ξ}^{0} and the javelin curve $\mathrm{G}_{\xi}(\Delta \mathrm{v})$. We have

$$\mathrm{EVPI}_{\xi}^{0} = \mathrm{E}_{\xi}\left[\mathrm{E}\left[\Delta \mathrm{V}_{a_0(\xi)} \mid \xi, \Pi_0 \setminus \xi_0\right]\right] = \sum_{a} \int_{\xi:\, a_0(\xi)=a} \mathrm{E}\left[\Delta \mathrm{V}_{a} \mid \xi, \Pi_0 \setminus \xi_0\right] \mathrm{dF}(\xi)$$

$$G_\xi(\Delta v) = P\left(E\left[\Delta V_{a_0(\xi)} \mid \xi, \Pi_0 \setminus \xi_0\right] > \Delta v\right) = \sum_a P\left(E\left[\Delta V_a \mid \xi, \Pi_0 \setminus \xi_0\right] > \Delta v, a_0(\xi) = a\right).$$

For single parameters ξ and familiar distributions $F(\xi)$, terms in these expressions may often be evaluated in closed form.

For many decision problems, the expected improvement $EVPI_\xi^0$ is equal to the information value $EVPI_\xi$ for the parameter ξ. Here

$$EVPI_\xi = E_\xi\left[E\left[\Delta V_{a^*(\xi)} \mid \xi\right]\right]$$

where $a^*(\xi)$ is the alternative a maximizing $E[V_a \mid \xi]$; $\Delta V_a = V_a - V_{a^*}$; and $a^* = a^*(\varnothing)$ is the overall optimal solution. Suppose, for example, that the function $E\left[\Delta V_a \mid \Pi\right]$ is multilinear in individual parameters $\xi \in \Pi$, that base values ξ_0 are equal to the means $E[\xi] = \bar{\xi}$, and parameters are probabilistically independent. Then because $a^*(\xi)$ maximizes $E[V_a \mid \xi]$, and

$$E[V_a \mid \xi] = E_{\Pi\setminus\xi}[E[V_a \mid \xi, \Pi \setminus \xi] \mid \xi] = E[V_a \mid \xi, \bar{\Pi} \setminus \bar{\xi}] = E[V_a \mid \xi, \Pi_0 \setminus \xi_0]$$

it follows that $a^*(\xi) = a_0(\xi)$. Here we have used independence and multilinearity. Again invoking independence and multilinearity, we obtain

$$\begin{aligned} EVPI_\xi = E_\xi\left[E\left[\Delta V_{a^*(\xi)} \mid \xi\right]\right] &= E_\xi\left[E\left[\Delta V_{a_0(\xi)} \mid \xi\right]\right] \\ &= E_{\Pi\setminus\xi}\left[E_\xi\left[E\left[\Delta V_{a_0(\xi)} \mid \xi, \Pi \setminus \xi\right] \mid \xi\right]\right] \\ &= E_\xi\left[E\left[\Delta V_{a_0} \mid \xi, \bar{\Pi} \setminus \bar{\xi}\right]\right] \\ &= E_\xi\left[E\left[\Delta V_{a_0(\xi)} \mid \xi, \Pi_0 \setminus \xi_0\right]\right] = EVPI_\xi^0 \end{aligned}$$

as claimed.

Endnotes

1. The exception is when there are only two alternatives. When there are only two alternatives and payoffs are multilinear in parameters ξ, we can show that sensitivity conclusions using a tornado diagram will always match those using a javelin diagram with independent uniformly distributed parameters having means equal to base values. This is beyond the scope of the present paper.