AUCTORES
Globalize your Research

International Journal of Clinical Case Reports and Reviews

Edouard Kujawski *

Open Access

Research Article

How Adjusting Elicited Health Utilities after the Fact can Adversely Affect Shared Decision Making

Edouard Kujawski

Independent researcher Berkeley, CA, USA.

***Corresponding Author:** Edouard Kujawski, Independent researcher Berkeley, CA, USA.

Received Date: August 15, 2025 | **Accepted Date:** September 02, 2025 | **Published Date:** September 12, 2025

Citation: Edouard Kujawski, (2025), How Adjusting Elicited Health Utilities after the Fact can Adversely Affect Shared Decision Making, *International Journal of Clinical Case Reports and Reviews*, 29(4); **DOI:**10.31579/2690-4861/949

Abstract:

Background: The elicitation of inconsistent health-state utility values (HSUVs) is a prevalent problem. There are two approaches to address this problem: (1) intervention during the elicitation process to ensure that patients estimate consistent HSUVs; (2) no intervention during the elicitation process and inconsistent HSUVs are adjusted after the fact. This paper studies three models recently proposed for adjusting inconsistent HSUVs and consistent HSUVs that some may consider unrealistic.

Analysis: The three models are analyzed using a sound theoretical framework: the mathematical equivalence of HSUVs elicited using the standard gamble and probabilities, the Fréchet bounds, and preference theory. It is proven that none of these models accounts for the Fréchet lower bound and health conditions that are preference substitutes.

Results: A clinical vignette proves these models may recommend treatments that result in premature death over treatments that cause acceptable adverse effects.

Conclusions: The three models are incorrect and may mislead patients and physicians to poor medical decisions. In the spirit of shared decision making, patients should be given the opportunity to reassess inconsistent HSUVs and confirm that the revised HSUVs reflect their preferences.

Key words: elicitation intervention; fréchet bounds; health-state utility values; preference complements; preference substitutes; reasonableness test

Introduction

The elicitation of health-state utility values (HSUVs) is daunting, and inconsistent elicited HSUVs are a prevalent problem. For instance, Dale et al.[1] found that at the individual level 41% of the elicited HSUVs for the joint health state (JHS) (incontinence & impotence) were larger than at least one of the HSUVs elicited for the constituent single health states (SHSs). This violates what they term *logical consistency* –i.e., rational people should not prefer a JHS to any of the SHSs. Logical consistency corresponds to the Fréchet [2] upper bound (FUB). It is also highly likely that the percentage of inconsistent HSUVs was higher than 41% because Dale et al. did not consider the Fréchet lower bound (FLB).

Two approaches to the problem of eliciting inconsistent HSUVs have been proposed:

1. Prevent inconsistent elicited HSUVs through interviewers intervening whenever necessary to ensure that patients estimate consistent HSUVs.[3]
2. No intervention during the elicitation process. Inconsistent elicited HSUVs are adjusted after the fact.[4,5]

The first approach requires trained and knowledgeable interviewers. The second approach requires realistic and mathematically valid models. However, the adjusted HSUVs may not accurately represent a patient's preferences.

Triantaphyllou and Yanase[4] (referred to as T-Y in this paper) proposed three models for adjusting inconsistent as well as consistent HSUVs which they say "may still not be realistic":

- Model (i): "Readjusting the original health utility values via an error minimization approach based on the monotonicity property."
- Model (ii): "Multiplicative functions for health states and a new model for adjusting the initial health-state utilities."
- Model (iii): "A combined approach for adjusting the initial health-state utilities."

This study has a sound theoretical framework: the mathematical equivalence of HSUVs elicited using the standard gamble (SG) and probabilities.[6] It uses probability theory and preference theory to prove that the T-Y models are incorrect. A simple clinical vignette (Text Box 1) demonstrates that these models can be misleading. Therefore, they are inappropriate for shared decision making (SDM) where reliable HSUVs are critical for patients and physicians to decide upon a preferred treatment.[7]

Text Box 1. Clinical vignette.

At their annual physical examination, patient HP is diagnosed to suffer from the asymptomatic health condition X. HP is otherwise in good health. HP is neither a trained nor innate probability assessor.

The physician is a proponent of SDM, and HP agrees to participate.

The physician informs HP that:

– X would reduce their expected years of life from 15 to 7, unless treated

– Treatment T_x has a 100% success rate in curing X

– T_x has a 100% probability of two side effects: a_i and b_j.

– HP agrees to partake in assessing the HSUVs for a_i, b_j and $a_i \& b_j$.

This paper proceeds as follows. The theoretical framework used for analyzing the T-Y models is presented. These models are analyzed, and it is demonstrated that they are inappropriate for life-critical SDM. Concluding remarks are presented.

Theoretical framework

Health-state utility values

Health states are identified by health conditions (HCs) (also termed "attributes" and "dimensions") and severity levels. Each combination of levels of HCs represents a unique health state. HSUVs are cardinal values specified on the (immediate death (ID) = 0.0, perfect health (PH) = 1.0) scale that measure the strength of a person's subjective preferences for health-related quality of life (HRQL).[8]

Estimating HSUVs is a challenging problem even for probability-savvy individuals. People are affected by the information that they receive about their medical conditions, emotional factors, and elicitation methods. HSUVs elicited by different methods may not agree and can affect treatment choices.[9] The SG has a theoretical foundation in von Neumann-Morgenstern expected utility theory [10], which establishes the validity of HSUVs elicited using the SG as a measure for HRQL.[11] Using arbitrary scales for HSUVs can lead to serious errors. [12, p.17]

Probabilities: Fréchet inequalities

Fréchet[2] proved that the joint probability of two events is bounded by the marginal probabilities of each event regardless of the dependence between them. For two events *A* and *B*,

$$\max(0.0, \mathrm{P}(A) + \mathrm{P}(B) - 1.0) \le \mathrm{P}(A \& B) \le \min(\mathrm{P}(A), \mathrm{P}(B)). \quad (1)$$

Few people realize they assign inconsistent values to joint probabilities. Osherson et al. [13] state: "It is striking to observe, for example, how few people realize that it is inconsistent to attribute probabilities of 0.8 to each of two sentences and probability 0.5 to their conjunction." From (1), the conjunction of 0.8 and 0.8 cannot be less than 0.6: $\mathrm{P}^{\mathrm{FLB}}(0.8 \wedge 0.8) = 0.8 + 0.8 - 1.0$.

Mathematical equivalence of HSUVs and probabilities

In the SG, an individual is asked to make the hypothetical choice between living for T years with health state a_i (health condition A with severity level i) and a gamble with a binary outcome (probability *p* of living in PH for T years or ID with probability $(1.0 - p)$). The probability *p* is varied until the individual is indifferent between living T years with a_i and the gamble. The indifference probability $p(a_i)$ corresponds to the individual's HSUV for a_i:[11] $\mathrm{U}(a_i) = p(a_i)$. The mathematical equivalence of HSUVs elicited using the SG and probabilities provides the basis for applying the power of probability theory to the problem of identifying inconsistent HSUVs.[6]

Consistent HSUVs: Fréchet inequalities

Given that HSUVs are mathematically equivalent to probabilities, the Fréchet inequalities play an important role in identifying inconsistent HSUVs. $\mathrm{U}(a_i \& b_j)$ is bounded by the FUB and FLB on conjunction irrespective of preference interactions: [6]

$$\max(0.0, \mathrm{U}(a_i) + \mathrm{U}(b_j) - 1.0) \le \mathrm{U}(a_i \& b_j) \le \min(\mathrm{U}(a_i), \mathrm{U}(b_j)). \quad (2a)$$

$$\mathrm{U}^{\mathrm{FUB}}(a_i \& b_j) = \min(\mathrm{U}(a_i), \mathrm{U}(b_j)). \quad (2b)$$

$$\mathrm{U}^{\mathrm{FLB}}(a_i \& b_j) = \max(0.0, \mathrm{U}(a_i) + \mathrm{U}(b_j) - 1.0). \quad (2c)$$

The FUB (2b) ensures logical consistency.[1] The FLB (2c) has significant implications for the disutility of multiple coexisting morbidities. The joint disutility cannot exceed ID or the sum of the individual disutilities:[6]

$$\bar{\mathrm{U}}^{\mathrm{FLB}}(a_i \& b_j) = \min(1.0, \bar{\mathrm{U}}(a_i) + \bar{\mathrm{U}}(b_j))$$

(3)

where $\bar{\mathrm{U}}(\cdot) \equiv 1.0 - \mathrm{U}(\cdot)$.

Preference interactions

HCs can be mutually utility independent (MUI), preference complements (PCs), or preference substitutes (PSs).[6] If a patient's preference for condition A is independent of the level of condition B and vice versa, A and B are said to be MUI: $\mathrm{U}^{\mathrm{MUI}}(a_i \& b_j) = \mathrm{U}(a_i) \times \mathrm{U}(b_j)$. If a patient believes that both A and B need to improve for their HRQL to improve, A and B are said to be PCs. PC HSUVs are positively correlated: $\mathrm{U}^{\mathrm{MUI}}(a_i \& b_j) < \mathrm{U}^{\mathrm{PC}}(a_i \& b_j) \le \mathrm{FUB}$. If a person believes that only A or only B needs to improve for their HRQL to improve, A and B are PSs. PS HSUVs are negatively correlated: $\mathrm{FLB} \le \mathrm{U}^{\mathrm{PS}}(a_i \& b_j) < \mathrm{U}^{\mathrm{MUI}}(a_i \& b_j)$.

Quality-adjusted life-years

The linear quality-adjusted life-year (simply termed *the QALY*) is presently the principal model for medical decision making (MDM). The expected number of QALYs for living y_i years in a health state x_i which has a probability of occurrence $p(x_i)$ is[11]

$$\mathrm{E\,Q}(p(x_i), x_i, y_i) = p(x_i)\times \mathrm{U}(x_i)\times y_i + (1.0 - p(x_i))\times y_i. \quad (4)$$

Analysis of T-Y models

In the following subsections, the T-Y models[4] are analyzed using the above theoretical framework and data shown in Table 1.

Model (i):[4] "Readjusting the original health utility values via an error minimization approach based on the monotonicity property."

The monotonicity property requires that a rational individual should not prefer a JHS to any of the constituent health states. Hence, Model (i) satisfies the FUB (2b). For consistency with probability theory, HSUVs elicited using the SG are also required to satisfy the FLB (2c). Model (i) does not address the FLB. For instance, it does not identify the HSUVs $\mathrm{U}(a_i) = 0.62,\ \mathrm{U}(b_j) = 0.73,\ \mathrm{U}(a_i \,\&\, b_j) = 0.15$ as inconsistent: $\mathrm{U}(a_i \,\&\, b_j) < \mathrm{U}^{\mathrm{FLB}}(0.62 \wedge 0.73) = 0.35\ (= 0.62 + 0.73 - 1.0).$

Model (ii):[4] "Multiplicative functions for health states and a new model for adjusting the initial health-state utilities."

Model (ii) posits $\mathrm{U}^{(\mathrm{ii})}(a_i \,\&\, b_j) = \mathrm{U}(a_i)\times \mathrm{U}(b_j).$ Keeney & Raiffa[14] proved that MUI is a necessary and sufficient condition for the multi-attribute utility function of *n* MUI attributes to be a multiplicative function of the single-attribute utility functions. They focused principally on decision making outside of the medical domain. They advocated assuming MUI with the significantly important qualifier[14, p. 244]: "the utility independence assumptions are appropriate in many realistic problems".

More recently, Howard and Abbas wrote[15, p. 578]

"We have several issues with this type of 'utility independence' reasoning…Enforcing these 'utility independence' assumptions result in functional forms that are simple, but quite frequently they will not represent the preference of the decision maker."

Experimental studies have concluded that the multiplicative model is not a suitable model for JHSUVs.[16] MUI is a strong assumption that is usually inappropriate for HSUVs.[17]

Model (iii):[4] "A combined approach for adjusting the initial health-state utilities"

Model (iii) posits that JHSs have a level of utility independence controlled by a parameter $0.0 \le \gamma \le 0.1$. Thus, the JHSUVs lie between the MUI HSUV and the FBU (2b) and they do not account for HCs that are PSs[6]. For instance, given $\mathrm{U}(a_i) = 0.62$ and $\mathrm{U}(b_j) = 0.73,$ Model (iii) predicts $0.45 \le \mathrm{U}^{(\mathrm{iii})}(a_i \,\&\, b_j) \le 0.62.$ This is wrong: HCs can be PSs, in which case $0.35 \le \mathrm{U}^{\mathrm{PS}}(a_i \,\&\, b_j) < 0.45.$

T-Y[4] recommend using Models (ii) and (iii) for JHSUVs that "would easily pass the previous monotonicity test but could still be considered as not realistic." This can mislead clinicians to recommend and patients to choose unwanted treatments. Case in point, a patient who wants to avoid treatments with HSUVs ≤ 0.45 chooses treatment T_X based on $\mathrm{U}^{(\mathrm{ii})}(a_i \,\&\, b_j)$ and $\mathrm{U}^{(\mathrm{iii})}(a_i \,\&\, b_j) \ge 0.45.$

Results and Discussion

The clinical vignette in Text Box 1 is analyzed assuming HSUVs that are elicited with and without intervention.

HSUVs elicited with intervention

Interviewers intervene when necessary to ensure that patient HP assesses consistent HSUVs which truly represent their preferences. Elicited single HSUVs (SHSUVs) are not always more correct than elicited JHSUVs.[1] HP adjusts the JHSUV and SHSUVs as shown in Text Box 1: $\mathrm{U}(a_i) = 0.55,\ \ \mathrm{U}(b_j) = 0.62,\ \ \mathrm{U}(a_i \,\&\, b_j) = 0.38.$ These HSUVs satisfy the FUB (= 0.55) and FLB (= 0.17).

HSUVs elicited without intervention

Interviewers do not intervene during the elicitation of HSUVs. The no-intervention elicited HSUVs shown in Table 1 violate the FLB (= 0.35). As discussed above, the T-Y models do not identify these HSUVs as inconsistent. T-Y recommend using Models (ii) and (iii) for JHSUVs that "could still be considered as not realistic."[4] These models predict the significantly different JHSUVs shown in Table 1.

Methods	$\mathrm{U}(a_i)$	$\mathrm{U}(b_j)$	$\mathrm{U}(a_i\&b_j)$
No intervention	0.62	0.73	0.15
Model (i)	0.62	0.73	0.15
Model (ii)	0.62	0.73	0.45
Model (iii)	0.62	0.73	$\in$ [0.45, 0.62]
$\in[\cdot,\cdot]$: lies in interval $[\cdot,\cdot]$			

Table 1: Elicited and adjusted HSUVs using T-Y models.

Decision analysis

For illustration, we consider the clinical vignette and data shown in Text Box 1 and Table 2, respectively. Patient HP has a complicated decision to make: "to be or not to be" treated with T_X? The expected number of QALYs for each alternative and set of HSUVs is calculated using (4). Table 2 summarizes the results and recommendations. The HSUVs elicited with and without intervention provide contradictory recommendations:

– HSUVs elicited with intervention. The prediction is: 15.0 YLs, 7.8 QALYs. The recommendation is "Yes T_x".

– HSUVs elicited without intervention and adjusted after the fact. Model (i) predicts 15 YLs and 2.25 QALYs. Models (ii) predicts 15 YLs and 6.75 QALYs. Based on the number of QALYs, Models (i,) and(ii) recommend "No T_x". Model (iii) recommends either "No T_x" or "Yes T_x" depending on the control parameter γ.

Kujawski et al.[18] proposed an intuitive reasonableness test that decision models used for SMD should pass to qualify as SDM tools: "Can a treatment that results in premature death trump a treatment that causes acceptable adverse effects?" A "Yes" answer may mislead clinicians into recommending and patients into choosing decisions with unintended consequences. As shown in Table 2, the three T-Y models fail this test.

Options	YLs	Elicitation	$P(a_i \& b_j)$	$U(a_i \& b_j)$	QALYs	RECs	RT
No T_x	7.00		1.00	1.00	7.00		
Yes T_x	15.00						
		Intervention	1.00	0.52	7.80	Yes T_x	P
		No intervention					
		Model (i)	1.00	0.15	2.25	No T_x	F
		Model (ii)	1.00	0.45	6.75	No T_x	F
		Model (iii)	1.00	$\in (0.45, 0.62]$	$\in [6.75, 9.30]$	?	F

REC: recommendation, RT: reasonableness test, P: pass, F: fail.

Table 2: Impact of HSUV elicitation and T-Y models on treatment recommendation.

Conclusions

The elicitation of reliable HSUVs is critical to ensure medical decisions that patients truly prefer. As shown in this paper, the three T-Y models[4,5] do not accurately account for individual preferences and the mathematical equivalence of HSUVs with probabilities elicited using the SG. Given consistent elicited HSUVs, it is not the function of clinicians to judge whether these are realistic or unrealistic.

A clinical vignette proves that the three T-Y models[4,5] may recommend treatments that result in premature death over treatments that cause acceptable adverse effects. This is a sure sign that these models are faulty and can be misleading. Well-trained interviewers are still essential to elicit reliable HSUVs. Practical tools are being developed to assist with the assessment of HSUVs, e.g., Gambler II.[19] The uncertainties of elicited HSUVs and calculated QALYs need to be addressed for sound SDM. Assuming point estimates causes false confidence in the analysis results.[20,21]

Disclosure of interest

The author reports no conflict of interest.

Funding

The author reports there is no funding associated with the work featured in this article.

ORCID

Edouard Kujawski https://orcid.org/0000-0002-6920-9078

References

1. W. Dale, P. Biller, A. Basu, A. Elstein, and D. Meltzer. (2011). The prevalence, correlates, and impact of logically inconsistent preferences in utility assessments for joint health states in prostate cancer. *Medical Care*, 49(1): 59–66.
2. M. Fréchet. (1935). Généralisations du théorème des probabilités totales. *Fundamenta Mathematicae*, 25(1): 379–387.
3. S. T. Stewart, L. Lenert, V. Bhatnagar, and R. M. Kaplan. (2005). Utilities for prostate cancer health states in men aged 60 and older. *Med Care* 43(4): 347-335.
4. E. Triantaphyllou and J. Yanase. (2020). How to identify and treat data inconsistencies when eliciting health-state utility values for patient-centered decision making. *Artif Intell Med*. 106, 101882.
5. E. Triantaphyllou and J. Yanase. (2021). Treatment selection for life-critical shared decision making under ranges of health-state utility scenarios. *Journal of Biomedical Informatics* 115, 1063604.
6. E. Kujawski. (2022). The importance of preference interactions in joint health-state utility values applied to decision analyses for shared decision-making. *ISSE* Transactions on Healthcare Systems Engineering.
7. Lafata, E. J., Brown, R. F., Pignone, M. P., Ratliff, S., & Shay, L. A. (2017). Primary care physicians' support of shared decision making for different cancer screening decisions. *Med Decis Making*, 37(1): 70-78.
8. G. H. Guyatt, D. H. Feeny, and D. L. Patrick. (1993). Measuring health-related quality of life, *Ann Intern Med*. 118(8): 622-629.
9. E. B. Elkin, M. E. Cowen, D. Cahill, M. Steffel, and M. W. Kattan. (2004). Preference assessment method affects decision-analytic recommendations: a prostate cancer treatment example. *Med Decis Making* 24(5): 504-510.
10. J. von Neumann and O. Morgenstern. (1947). Theory of games and economic behavior. *Princeton University Press*.
11. M. G. M. Hunink, M. C. Weinstein, E. Wittenberg, M. F. Drummond, J. S. Pliskin, et al., (2014). Decision making in health and medicine: integrating evidence and values. *Cambridge University Press*, 2nd edition.
12. A. E. Abbas. (2018). Foundations of multiattribute utility. *Cambridge University Press*.
13. D. Osherson, D. Lane, P. Hartley, and R. R. Batsell. (2001). Coherent probability from incoherent judgment. *J Exp Psychol Appl*. 7(1): 3-12.
14. R. L Keeney and H. Raiffa. (1976). Decisions with multiple objectives: preferences and value tradeoffs. John *Wiley* & Sons.
15. R. A. Howard and A. E. Abbas. (2015). Foundations of decision analysis. Pearson.
16. A. Z. Fu and M. W. Kattan. (2008). Utilities should not be multiplied: evidence from the preference-based scores in the United States. *Med Care*, 46(9): 984-990.
17. E. Kujawski. (2019). On the appropriateness/inappropriateness of the Keeney-Raiffa multiplicative utility function for medical decision making. PPI SyEN, 79: 5–17.
18. E. Kujawski, E. Triantaphyllou, and J. Yanase. (2019). Additive multicriteria decision analysis models: misleading aids for life-critical shared decision making. *Med Decis Making* 39(4): 437-449.
19. A. C. Behney. (2020). Ignoring uncertainty in predictor variables leads to false confidence in results: a case study of duck habitat use. *Ecosphere* 11(10): 1-13.
20. A. A. Adejare Jr. and M. H. Eckman. (2020). Automated tool for health utility assessments: the Gambler II. *MDM Policy & Practice*, 5(1): 1–12.
21. D. J. Spiegelhalter, R. C Franklin, and K. Bull. (1989). Assessment, criticism and improvement of imprecise subjective probabilities for a medical expert system. Proceedings of the fifth workshop on uncertainty in artificial intelligence: 335-342.

DOI:10.31579/2690-4861/949