Perspectives on cost-effectiveness thresholds in the United States

Moderated by:

- Dr. Steven Pearson, President
- Dr. Rick Chapman, Director of Health Economics



© Institute for Clinical and Economic Review, 2019

Webinar 3: Willingness to pay Part 2: Using past funding decisions, value of a statistical life, and relative risk aversion as the basis for determining a cost-effectiveness threshold in the US

Main Presentations:



Sean Sullivan University of WA School of Pharmacy



Chuck Phelps, PhD University of Rochester



Value of Statistical Life and Observed Resource Allocation Decisions as Approaches to CEA Threshold Development

> Sean D. Sullivan, PhD Professor and Dean University of Washington



Value of Statistical Life

 Issues pertaining to the value of life (statistical or otherwise) are amongst the most sensitive and controversial in economics.

- However, VSL has been an integral part of benefits assessment in the US for 30+ years
 - Department of Transportation and EPA, in particular.



Value of Statistical Life

- The economic measure for the optimal deterrence amount is the risk-money tradeoff for very small risks of death.
- Since the concern is with small probabilities, not the certainty of death, these values are referred to as the value of statistical life (VSL).
- Economic estimates of the VSL amounts have included evidence from market decisions that reveal the implicit values reflected in behavior as well as the use of survey approaches to elicit these money-risk tradeoffs directly.
- Government regulators in turn have used these VSL estimates to value the benefits associated with risk reduction policies.

Value of Statistical Life

- VSL serves as a measure of the deterrence amount for the value to the individual at risk of preventing accidents and as a reference point for the amount the government should spend to prevent small statistical risks.
- Because the concern is with <u>statistical lives</u>, not identified lives, analyses of government regulations now use these VSL levels to monetize risk reduction benefits.



Value of Statistical Life Varies by Age



Viscusi 2010, ASHE

Value of a Statistical Life-Year



Viscusi 2010, ASHE

\$/QALY based on VSL

VOL 20/NO 3, JUL-SEP 2000

Willingness to Pay for a QALY • 339

Data Source#	Turne of Bohovier/Diels Applyroot	Value of Life	¢/0.41.V
	Type of Benavior/Risk Analyzed†	(1997\$, 0.5.)	\$/QALY
Questionnaire, 1994	Road accidents & safety	\$1,230,828	58,639
Montreal-based survey, 16 random firms w/100+ employees	Car safety—willingness to have an air-bag system in their car	\$2,310,012	104,796
Questionnaire, 1982	Tradeoffs among fatal and nonfatal risks	\$2,301,982	109,670
Questionnaire, 1992	Traffic safety	\$2,670,436	127,224
Mail questionnaire 1984	Job fatality risk, willingness to pay	\$4,101,153	195,386
Questionnaire, Pittsburgh, PA, 1986	Automobile safety	\$9,029,041	380,491
Mail questionnaire 1984	Job fatality risk, willingness to accept	\$10,514,986	500,952
Montreal-based survey, 16 random firms w/100+ employees	Subjective valuation of on job risk	\$25,926,349	1,176,171

Administration; NHDS = National Hospital Discharge Survey; NHIS = National Health Interview Survey; NMCUES = National Medical Care Utilization USDT = U.S. Department of Transportation.

- UNIVERSITY of WASHINGTON -

Source (dollar year)	VSL estimate (range)	GNI per capita	Ratio of VSL to GNI per capita	WTP for 1 in 10,000 risk change	WTP as percent of GNI per capita
USEPA (2010a) (2006 USD)	\$7.4 million (± \$4.7 million)	\$47,390	156 (57, 255)	\$740 (\$270,\$1210)	1.6% (0.6%, 2.6%)
USDHHS (2016) (2014 USD)	\$9.3 million (\$4.4 million, \$14.2 million)	\$56,160	166 (78, 253)	\$930 (\$440,\$1420)	1.7% (0.8%, 2.5%)
USDOT (2016) (2015 USD)	\$9.6 million (\$5.4 million, \$13.4 million)	\$57,900	166 (93, 231)	\$960 (\$540,\$1340)	1.7% (0.9%, 2.3%)
OECD (2012) (2005 USD)	\$3.0 million (\$1.5 million, \$4.5 million)	\$30,601	98 (49, 147)	\$300 (\$150, \$450)	1.0% (0.5%, 1.5%)

US VSL Estimates By Agency*

• Summary –

- There is much heterogeneity in VSL estimates as much to do with methods, data sources, questions as expressions of WTA risk.
- \$/QALY from VSL estimates produce a range \$100,000/QALY \$300,000/QALY
- In actual social choices, once the individual is identified (known), there is no amount of \$ that can be offered that can compensate for that death (John Broome, 1978).
- There is no precise, single decision rule that can be derived from this literature for use in a US health care context.

*Robinson, et al 2019, JCBA

-UNIVERSITY of WASHINGTON-

Do Actual Resource Allocation Decisions Provide a Framework to Support a US CEA Threshold?

- Evidence from actual decisions that individuals and organizations make is potentially more informative than tradeoffs based on hypothetical situations if suitable market data exists.
 - The risks, benefits and costs to organizations are real, so they do not have to engage in a WTP or WTA thought experiment.
 - Urban myth that the \$50,000 QALY threshold in the US is linked to the <u>political decision</u> to expand Medicare to include renal dialysis. Real value of dialysis today is above \$150,000/QALY.
- There is a small literature on this in the U.S.
 - Braithwaite et al, 2008, MDM
 - Chambers et al, 2010, MDM

Braithwaite, et al approach

- To investigate whether the \$50,000 per QALY rule is consistent with current resource allocation decisions.
- Two approaches:
 - To infer a lower bound, estimated the incremental cost-effectiveness of recent (2003) versus pre-"modern era" (1950) medical care in the United States.
 - To infer an upper bound, estimated the incremental cost-effectiveness of unsubsidized health insurance versus self-pay for nonelderly adults (ages 21-64) without health insurance.
- Conclusion:
 - "The \$50,000 per QALY threshold is inconsistent with observed spending behavior in the United States at the societal level."

-UNIVERSITY of WASHINGTON-

TABLE 4. Incremental Cost-Effectiveness From (A) Modern Health Care Over Lifetime of US Birth Cohort and (B) Health Insurance Over 1 Year for a Cohort of Nonelderly Adults

A. Incremental Cost-Effectiveness From Modern Health Care

	Lifetime Care Costs, 1950 Medical Care	Life Expectancy, 1950 Medical Care	Lifetime Care Costs, 2003 Medical Care	Life Expectancy, 2003 Medical Care	Incremental Cost-Effectiveness, 2003 Health Care Compared With 1950 Health Care
Base case	\$16,600	29.70 LY	\$135,000	30.35 LY	\$183,000/LY
Health benefit accrues at earliest possible age	\$16,700	29.48 LY	\$135,000	30.35 LY	\$137,000/LY
Health benefit accrues at latest possible age	\$16,500	29.88 LY	\$135,000	30.35 LY	\$253,000/LY
Analysis excludes children	\$18,700	19.45 LY	\$169,000	21.02 LY	\$ 95,000/LY
Incorporates IOM estimate	NA	NA	NA	NA	NA
Analysis does not discount	\$56,700	72.60 LY	\$509,000	77.30 LY	\$ 96,000/LY
Optimistic assumption, QALYs conferred	\$16,600	28.68 QALY	\$135,000	29.77 QALY	\$109,000/QALY
Pessimistic assumption, QALYs conferred	\$16,600	29.19 QALY	\$135,000	29.77 QALY	\$204,000/QALY

B. Incremental Cost-Effectiveness From Health Insurance

	Costs, No Health Insurance	Life Expectancy, No Health Insurance	Costs, Health Insurance	Life Expectancy, Health Insurance	Incremental Cost-Effectiveness, 2003 Health Insurance
Base case	\$443	21.006 LY	\$3826	21.019 LY	\$264,000/LY
Health benefit accrues at earliest possible age	NA	NA	NA	NA	NA
Health benefit accrues at latest possible age	NA	NA	NA	NA	NA
Analysis excludes children	NA	NA	NA	NA	NA
Incorporating IOM estimate	\$443	21.000 LY	\$3826	21.019 LY	\$178,000/LY
Analysis does not discount	\$443	34.339 LY	\$3826	34.360 LY	\$182,000/LY
Optimistic assumption, QALYs conferred	\$443	18.688 QALY	\$3826	18.718 QALY	\$113,000/QALY
Pessimistic assumption, QALYs conferred	\$443	18.707 QALY	\$3826	18.718 QALY	\$297,000/QALY

We argue that the incremental cost-effectiveness of modern health care is a plausible lower bound for society's willingness to pay (WTP) for health care, and the incremental cost-effectiveness of health insurance for those without employer- or government-subsidized insurance is a plausible upper bound. Our base case analyses suggest a range of \$183,000/LY to \$264,000/LY for society's WTP, whereas our sensitivity analyses suggest broader ranges (\$95,000/LY to \$264,000/LY and \$109,000/QALY to \$297,000/QALY). Individuals without health insurance are assumed to forgo the proportion of health benefits (35%) that arise from having insurance. All costs are inflated to 2003 US\$. LY indicates life-year; OALY, quality-adjusted life-year; IOM, Institute of Medicine, NA, not applicable.

14

Chambers, et al approach

- The assess whether an implicit cost-effectiveness threshold exists in the Medicare program using past funding decisions (NCDs) as the source of evidence.
 - 64 coverage decisions were linked with cost-effectiveness evidence

Cost-Effectiveness Estimate	Frequency Positive Coverage Decisions	Noncoverage Decisions
Dominant	20	3
< \$50,000 per QALY	12	6
\$50,000-\$100,000 per QALY	8	0
> \$100,000 per QALY	9	1
Dominated	0	5

 Table 2
 Frequency of Identified Cost-Effectiveness Studies

QALY, quality-adjusted life year.

- Medicare NCDs follow a "loose" cost-effectiveness logic, but reflect a broader set of decision factors.
- ".....no clear evidence of a CEA threshold."

A Few Concluding Comments

- The US Health Care System is pluralistic, heterogeneous and comprised of many public and private purchasers.
- ICER is not responsible for a health care budget and does not make purchasing decisions.
- Perhaps ICER should not have a single CEA threshold at all? Rather, the final HTA reports might display various levels of pricing across a range of thresholds from \$10,000/QALY to \$1,000,000/QALY.
 - Let the budget holders apply the evidence of CER and CEA in a manner that is consistent with their own regulations and market conditions.

A New Method to Determine the Optimal Cost-Effectiveness Cutoff

Charles E Phelps, PhD University Professor and Provost Emeritus University of Rochester

For presentation at ICER Webinar, July 26, 2019

Previous methods to determine "the" cutoff

League tables

- Arbitrary choices of "reasonable"
- No certainty that previous choices made with cost-effectiveness in mind
- In US, bias towards over-use from tax exemption for employer-sponsored health insurance (mimicked in Medicare)
- Value of Statistical Lives
 - Conflates WTP to avoid risk and WTA to accept risk
 - 5X bias in direct comparisons in health care
 - How to "spread" the VSL over life years remains unresolved
 - Hirth et al. find 17X range of estimates over 42 studies
- Direct WTP estimates
 - Review of 383 European studies gives 400X to 1200X range of estimates
- Hall-Jones cost of producing life-years
 - Not the same as WTP, also biased by US health insurance structure

• ALL JUST DIFFERENT "CALIBRATION" METHODS

Consensus numbers

- Neumann, Cohen and and Weinstein (2014) \$100K to \$150K
 - About 2X to 3X per capita GDP in US
- WHO (2003) 1X to 3X per capita GDP
 - Now backing away from fixed numbers
- ICER (current web site) \$100K to \$150K
- British NHS £20,000 to £30,000
 - About 0.7X to 1X per capita GDP
 - Operationally, may be £12,000 or so
 - Allows larger values for rare diseases, end of life,
 - UK Dept. of Treasury recommends £60,000 (2X per capita GDP)
- Sweden "rule of thumb" *Kr* 500,000 ≈ 1X per capita GDP
 - Commonly exceeded for serious diseases

A New Method: General Concepts

For M = money income and K = optimal cutoff:

Optimal K = U(M)/U'(M)

Garber and Phelps, 1997, *Journal of Health Economics* Pertains to a single utility-maximizing rational individual

From that K/M = 1/E(U,M)

Intuition: E(U,M) represents the **opportunity cost of consumption** (in utility). Small E(U,M) means that you are not giving up much utility as you expand medical care use (and conversely).

How should K change with M?

For *all* utility functions:

 $E(K,M) = r^* + E(U,M) > 0$ where $r^* =$ relative risk aversion measure

If E(K,M) > 1 then E(K/M, M) > 0 (this is an empirical issue)

Model is calibrated using estimates of relative risk aversion (r*)

- Absolute risk version: r = -U''(M)/U'(M)
- Relative risk aversion r* = rM = -(income elasticity of U'(M))
 - DARA = Decreasing absolute risk aversion
 - CARA = Constant absolute risk aversion
 - IARA = Increasing absolute risk aversion
 - DRRA = Decreasing relative risk aversion
 - CRRA = Constant relative risk aversion
 - IRRA = Increasing relative risk aversion
- Arrow (1971) argued for DARA and IRRA
- General agreement for DARA, less so for IRRA

Behavior of absolute risk aversion

Behavior of relative risk aversion

Finding a Useful Functional Form

- Most previous work (Feldstein and Friedman, Arrow, Garber and Phelps, Murphy and Topel, Becker et al, Cordoba and Ripoll, others) has used either CARA or CRRA utility functions
 - Convenient mathematically but highly restrictive
- Need a general functional form that can be identified with observable data (e.g., relative risk aversion parameter)
- The Weibull function serves well

•
$$U(M) = 1 - e^{-\gamma M^{C}} = \frac{\left[e^{\gamma M^{C}} - 1\right]}{e^{\gamma M^{C}}} \rightarrow U$$
 is [0,1] for M [0, ∞]

•
$$r^* = C\gamma M^C + (1-C)$$

Calibrating C and y

- CARA when C = 1
 - DARA for C < 1
- "Almost" CRRA when C = 0.5 or less
 - IRRA for $\gamma > 0$
 - For wide range of M, r* almost invariant for $C \le 0.5$
- For given value of C $(0.5 \le C \le 1)$
 - Calculate γ for chosen r* at M = \$50,000
 - Calculate K and K/M for various M, given C and γ .
- Find out what it shows!

How big is r*?

- Labor supply
 - Chetty summarizes dozens of labor supply results: r* = 0.7 to 1.0
 - Rules out anything larger than r* = 2
- Dutch panel : r* = 0.93
- Macro simulations r* "about" 0.8
- Self-reports of "happiness" 0.88 in 23 developed countries
 - (Gallup poll data)
- I use r* = 1 in simulations
 - r* = 0.8 to 0.9 probably closer to truth
 - Slightly inflates estimates of K and K/M

Results: Optimal K



Results: Optimal K/M



Key findings

- At \$50,000 income and "central" parameters, optimal K/M ≈ 2
- Optimal K and K/M rise with income
 - For all parameter values assessed
 - Consistent with medical care being a "luxury good"
 - Consistent with large employers offering array of benefit options to employees
- Remember, it's all driven by the opportunity cost of consumption
- Major extrapolation beyond calibration values probably unwise
 - Too heavily driven by functional form that has only 2 parameters
- Findings modestly sensitive to chosen value of r*
- Findings more sensitive to chosen value of C
 - Need better assessment of DRRA vs. IRRA vs. CARA vs. CRRA vs....

Ironic result: Garber and Phelps had "WAWR"

- Garber and Phelps used exponential function (C = 1 in Weibull)
 - Ruled out by evidence that utility has DARA
 - C=1 gives overly low estimates of K/M
- We also said "r* is about 2.0"
 - Far too high by current evidence
 - Optimal K/M rises with r*
- The errors offset
- In the Weibull model, for C=1 and $r^* = 1.7$

K/M ≈ 2

Moving from single person to societal perspective for choosing K or K/M

- Question 1: single or multiple cutoffs?
 - Medicaid vs. Medicare vs. private plans
 - Plans can compete on generosity
 - Scope of benefits
 - Prior authorization stringency
 - Breadth of panel and out of panel benefits
- Question 2: Can individuals supplement "central" plan care?
 - "Harley Street" medicine (5% of health spending) and EEU "tourism"
 - Insured or not? (UK allows insurance; Canada bans it)
- Question 3: Median voter vs. welfare maximizing cutoff?
 - Skewed income distribution assures that median voter model will choose smaller K than welfare maximizing model in almost all settings

Thank you for your attention



Pop Quiz: what's your E(U,M)?

If your permanent real buying power increased (decreased) by 10% and you adjusted all of your spending to accommodate that change, how much would your happiness increase (decrease)?

If "5%," then E(U,M) = .5, so K/M = 2

If "1%" then E(U,M) = 0.1, so K/M = 10

If "9%" then E(U,M) is 0.9, so K/M = 1.11...

How does this work???



Discussion

Responders: Chris McCabe, Jason Wasfy

Comments on "A new method to determine the optimal cost effectiveness cut-off"

INSTITUTE OF
HEALTH ECONOMICSDr. Christopher McCabeAlberta CANADAJuly 26, 2019ICER Webinar Series

STITUTE OF EALTH ECONOMICS QUESTIONS RE: MODE

- The model, like all good economic models, is based upon estimating risk aversion for marginal changes. However, most health care expenditures for most households would represent catastrophic expenditures in the absence of health insurance. The evidence on risk attitude to extreme events is that it is different.
 - What would be the impact of incorporating catastrophic risk aversion in to the model on the threshold?
- The model may be missing an important level of decision making for health insurance i.e. the household. The risk aversion of household decision making may be different from individuals; especially for households with children in. Parents are likely much more risk averse over their children's health than their own.
 - How would adopting a household risk aversion, rather than an individual risk aversion impact upon the threshold?
- The model is explicitly about opportunity cost in terms of individual utility of consumption. The model
 doesn't capture any social welfare from the availability of health care to fellow citizens. Whilst the US is
 likely the most individualistic of the OECD nations, Medicare, Medicaid and the legal right to 'emergency
 care' all indicate that there is a degree of social solidarity for access to health care.
 - Is this model relevant to coverage decisions for social health care payers/providers such as Medicare, Medicaid and arguably even Veteran's Affairs?

 <u>Median voter</u>: If individual values were aggregated under a median voter model; this would lead to those individuals with a higher value of k being denied access to some care that they would be willing to pay for.

HE

NOMICS

– Is this a threat to the financial sustainability of the risk pool if these people can opt-out?

Questions re: scaling from individual to society

- Would top up insurance combined with a mandated 'social insurance' be a mechanism for addressing this?
- 5) The 1997 paper this builds upon, emphasizes that k varies not just by risk attitude, but by age, gender & income. Aggregating up from individuals would need to be done in a way that respects these characteristics.
 - As different K translate into different insurance premia does this become de facto price discrimination and is that sustainable?

Next webinar:

Tue, Jul 30, 2019 12:00 PM - 1:00 PM EDT

Webinar 4: Opportunity cost as a basis for a cost-effectiveness threshold in the United States

How can this concept be measured in the US context? What are the specific strengths and weaknesses of using this approach to define a corresponding cost-effectiveness threshold in the US?

Hosts: Steve Pearson and Rick Chapman Lead Presentation: Karl Claxton Responders: Sean Sullivan, Lou Garrison