Sampling Uncertainty and Cost-Effectiveness Analysis (Part 1a)

Henry A. Glick, Ph.D.

EPI 550

April 17, 2020

Confidence About Value for the Cost

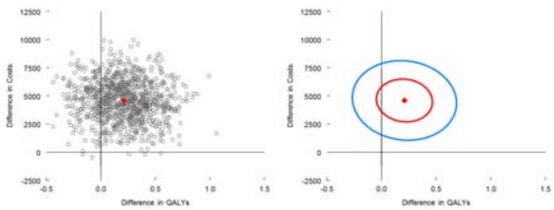
- Common goal of economic analysis: identify when we can be confident that a therapy is good value compared to another
- Threat to confidence: economic result observed in experiment may not reflect result in the population
  – Single sample drawn from population
- Referred to as sampling (or stochastic) uncertainty
- Methods for estimating sampling uncertainty for economic outcomes have much in common with methods used for clinical findings
  – But there are ways that they differ

Outline

- Familiarize you with methods used in literature to identify when we can and cannot be confident about a therapy's value
  – Cost-effectiveness plane
  – Acceptability curves
  – CI for ICER
  – CI for NMB
  – Value of information
- Goal: demonstrate quantification and interpretation of sampling uncertainty using these methods
  – Including where people have gone wrong
- Don't focus on technical aspects of estimation

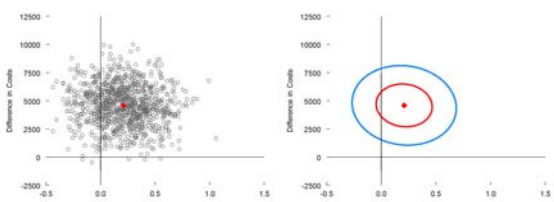
Cost-Effectiveness Plane

Joint Distribution of Cost and Effect



- Bootstrap of patient level data (left)
- Second order Monte Carlo (decision analysis with variables represented by distributions) (left)
- Bivariate normal curves (Δc, SEc, Δq, SEq, ρ) (right)

Estimates From the 2 Cost-Effectiveness Planes

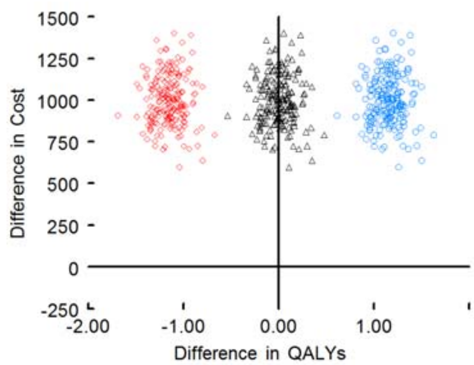


- Mean cost difference, $4600, SE, 1803
- Mean QALY difference, 0.2090, SE, 0.2430
- Correlation of difference, -0.045
- ICER Point estimate = 22,010 (4600 / 0.2090)

Information from Plane

- Cost-effectiveness plane provides information about point estimates, confidence intervals and p-values for:
  - Difference in effect
  - Difference in cost
  - Cost-effectiveness analysis

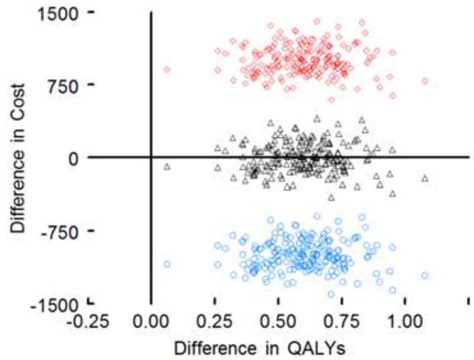
In Which Experiment(s) is ΔQ Significant?



Red and blue
(because all of their densities fall on one side of 0 on Y-axis)

Black triangles not significantly different (because too large a density falls on each side of 0 on Y-axis)

In Which Experiment(s) is ΔC Significant?

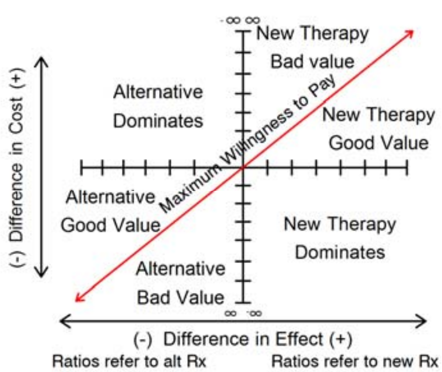


Overview1.tc

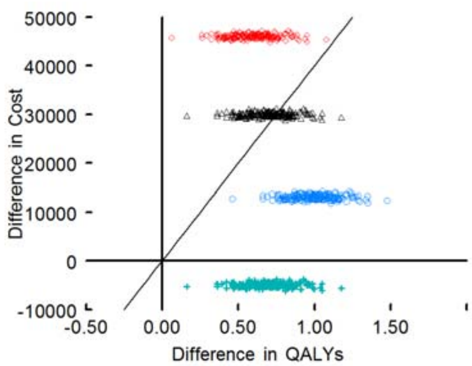
Red and blue
(because all of their densities fall on one side of 0 on X-axis)

Black triangles not significantly different
(because too large a density falls on each side of 0 on X axis)

Value and the Cost-Effectiveness Plane



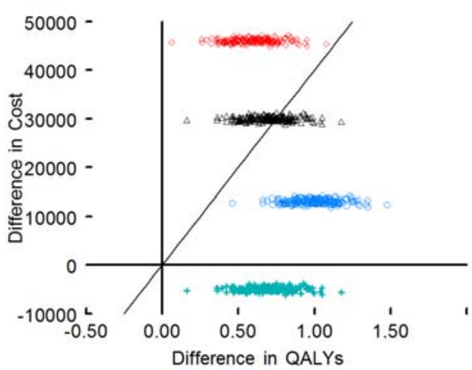
In Which Experiments Can We Be Confident of Value?



Red, blue, and cyan
(because all of their densities fall on one side of WTP)

Black triangles not confident because large fractions of density fall on both sides of WTP

Confidence Statements for Red, Blue, and Cyan?



For cyan, confident of dominance

(because all of density falls below X-axis and to right of Y-axis)

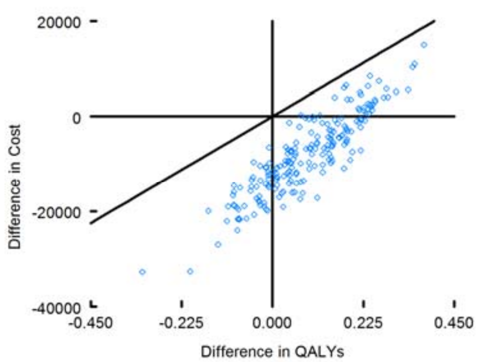
For blue, confident of good value

(because all of density falls above x axis, to right of y axis, and below WTP line

For red, confident of bad value

(because all of density falls above x axis, to right of y axis, and above WTP line)

What Can We Conclude About ΔC, ΔQ, Value?



Can't be confident about difference in cost

(because too large a density above and below X-axis)

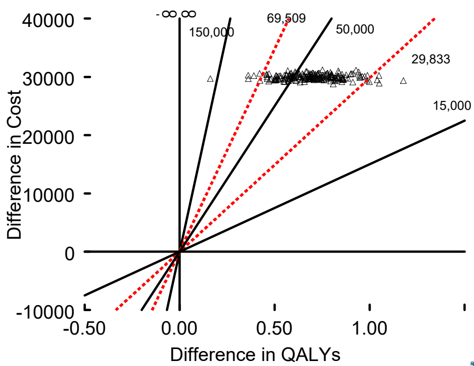
Can't be confident about difference in effect

(because too large a density to the left and right of Y-axis)

Can be confident of cost-effectiveness

(because all density below WTP line)

What if there is No Single Recognized



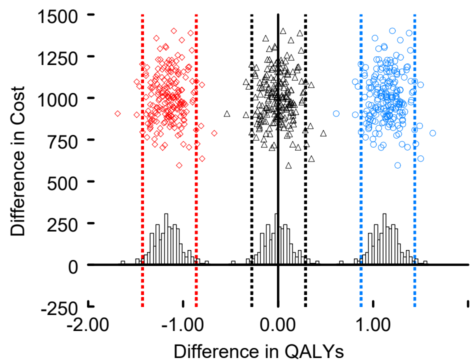
Provide a statistic that allows decision makers to determine if – based on their own WTP – they can be confident of value

---

Confidence Intervals

- Graphs above provide examples of 0 (for differences in means, including NMB), 1 (for OR and RR), or willingness to pay (W) (for CI for CER) falling either well inside or fully outside distribution of results
- Don't typically require that results be fully outside distribution to conclude they significantly differ from 0, 1, or W
  - Parametrically never happens
- Usual strategy: Identify a tolerance – e.g., 2.5% for 95% confidence or 5% for 90% confidence – for the maximum fraction of results that can fall on one side of 0, 1, or W
- Conclude with 95% confidence that result excludes 0, 1, or W if 0, 1, or W fall outside 95% CI

95% CI, ΔQ

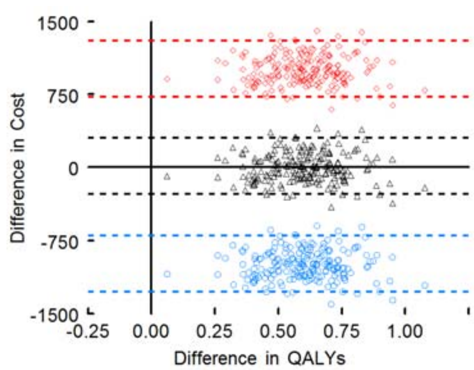


Overview2.histo&ci.tc

Can be 95% confident of a difference for red and blue (because 0 on X-axis does not fall within the 95% CI)

Can't be 95% confident of difference for black triangles (because 0 on X-axis falls within 95% CI)

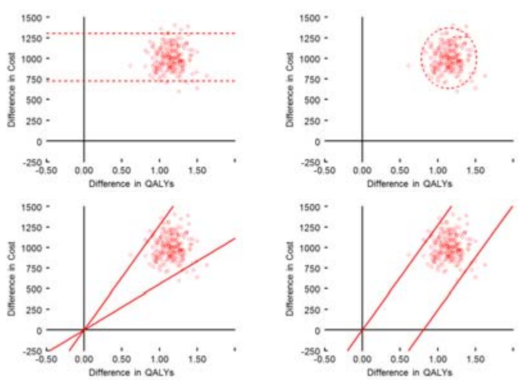
95% CI, ΔC



Can be 95% confident of a difference for red and blue (because 0 on Y-axis does not fall within the 95% CI)

Can't be 95% confident of difference for black triangles (because 0 on Y-axis falls within 95% CI)

95% CI for ICER?



95% CI

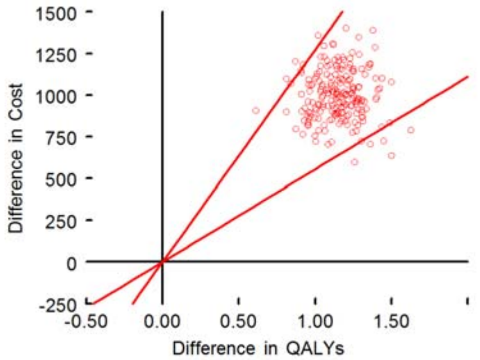
Upper left: CI for ΔC

Upper right: 95% confidence ellipse (CE) around the point on the C/E plane defined by ΔC and Δq (CE for point, not CI for ICER)

Lower right: CI for NMB

Lower left: 95% CI for the ICER

For Which Values of W Can We Be Confident?



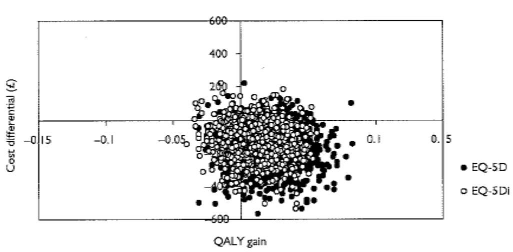
For Which Values of W…?

Confident of bad value for values of W between 0 (X-axis) and lower right line

Confident of good value for values of W between upper left line and ∞ (Y-axis)

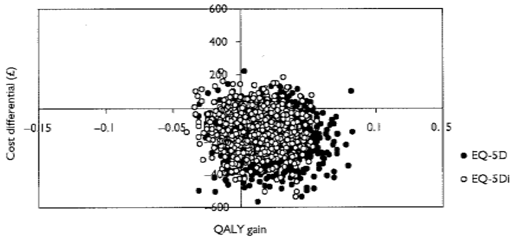
Not confident therapies differ in value for values of W between the lower right line and the upper left line

Conclusions From Cost-Effectiveness Plane??



Cochrane T, Randomized controlled trial of the cost-effectiveness of water-based therapy for limb osteoarthritis.Health Technology Assessment. 2005; 9. Figure 14

Cost-Effectiveness Plane??

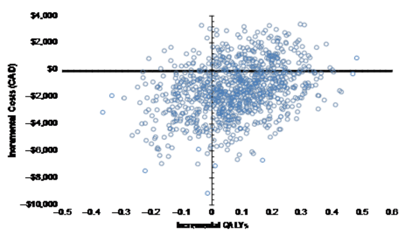


Reported cost difference:	-134.1, 95% CI -141 to -127.2 (p<0.05)
Reported EQ-5Di difference:	0.013, 95% CI, 0.0117 to 0.0138 (p<0.05)
Reported ICER:	5008, 95% CI, -22314 to 25288/Q

Probable mistake: mistook bootstrap SE for bootstrap SD and divided by $N^{0.5}$

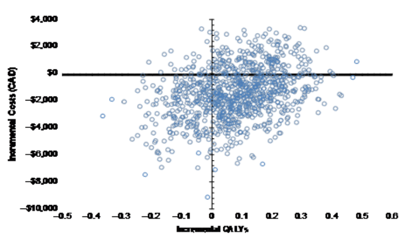
Cochrane T, Randomized controlled trial of the cost-effectiveness of water-based therapy for limb osteoarthritis.Health Technology Assessment. 2005; 9. Table 25.

Conclusions From Cost-Effectiveness Plane?



Brown ST, et al. Cost-effectiveness of insulin glargine versus sitagliptin in insulin-naïve patients w/ T2DM. Clin Therapuetics.2014; 36: 1576-87

Cost-Effectiveness Plane



Reported cost difference:	-1434, 95% CI -1553 to -1314
Reported EQ-5Di difference:	0.076, 95% CI, 0.068 to 0.084
Reported ICER:	Dominant, 95% CI, Dominates to $3179/Q

Probable mistake: mistook bootstrap SE for bootstrap SD and divided by $N^{0.5}$

Brown ST, et al. Cost-effectiveness of insulin glargine versus sitagliptin in insulin-naïve patients w/ T2DM. Clin Therapuetics.2014; 36: 1576-87. (Tables IV & V)

Sampling Uncertainty Issues

- # of methods available for expressing uncertainty
  - Acceptability curve
  - CI for ICER
  - CI for NMB
  - Value of information
- What is threshold, maximum willingness to pay?
  - Differs across jurisdictions
  - Differs within jurisdictions
- Should we be 95% confident?
  - A lot of economists claim not

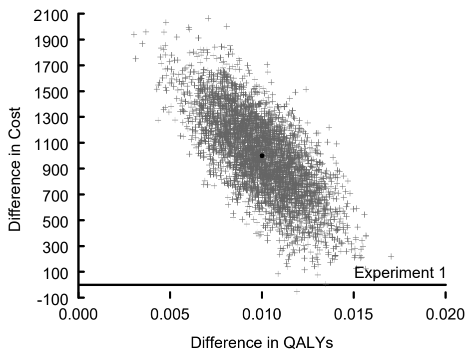
First Example:

(Nonparametrically) all replicates on one side of Y-axis and naïve ordering works (easiest case)

Experiment 1

- Therapy A vs Therapy B (A – B)
- Δcost = 1000 (SE: 324.9, p=0.002)
- ΔQALYs = 0.01 (SE: 0.001925, p=0.000)
- A is significantly more costly and significantly more effective
  - CER = 1000 / 0.01 = 100,000 / QALY gained
- 250 participants in each arm of the trial
- Correlation between difference in cost and effect is -0.71015

Distribution of Results



Acceptability Curve

"Counting" Method 1: Acceptability Curve

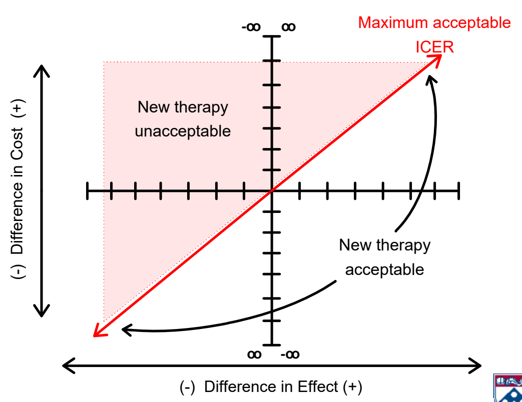
- Previously said usually identify a tolerance – e.g., 2.5% for 95% confidence – for the maximum fraction of results that can fall on one side of 0, 1, or W
- Can determine fraction that falls on one side of W by counting / estimating density of results distribution falling on each side of W
- Referred to as acceptability curve

Parametric or Nonparametric Construction

- Can be constructed nonparametrically or parametrically
- Nonparametric construction usually derived by counting bootstrap/second order Monte Carlo replicates
  - Does not assume bivariate normality
  - Particularly for acceptability curve and CI for CER, calculating fraction falling on each side of exact same lines through origin
- Parametric construction generally based on (rearrangement of) Fieller's theorem formula for CI for CER (i.e., transformation of same formula)
  - Assumes difference in costs and effects distributed bivariate normal

Acceptability Curve

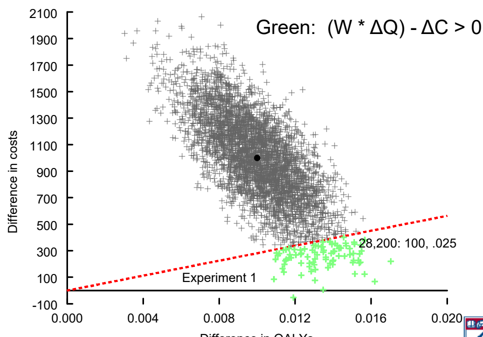
- Acceptability criterion defined on cost-effectiveness plane as a line passing through origin with slope equal to WTP
- When applied to positive values of WTP, proportion of distribution of difference in cost and effect below and to right of line is "acceptable" (i.e., has positive NMB)
  - Proportion acceptable for one therapy = proportion unacceptable for alternative therapy
- Proportion above and to left of line is "unacceptable“
  - Proportion unacceptable for one therapy = proportion acceptable for alternative therapy



Counting Methods

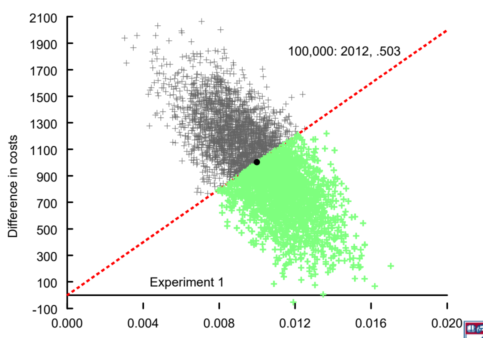
- When all replicates are on one side of Y-axis, one can:
  - Calculate ratios and count ratios < WTP
- Whether or not all replicates are on one side of the Y-axis, one can:
  - Calculate NMB using WTP; count values of NMB > 0
  - Calculate MB for each Rx using WTP; count replicates for which Rx A's MB greater than Rx B's

% Acceptable, W = 28,200



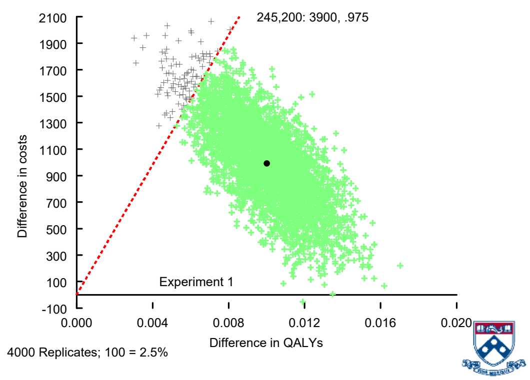
4000 Replicates; 100 = 2.5%

% Acceptable, W = 100,000

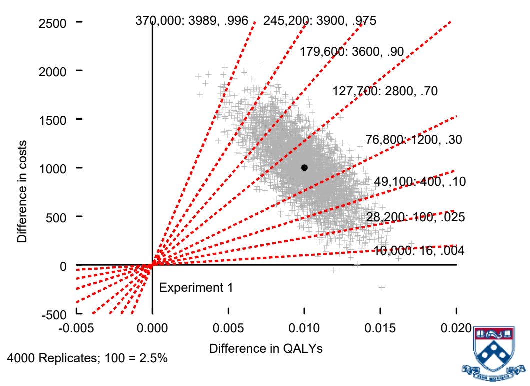


4000 Replicates; 100 = 2.5%

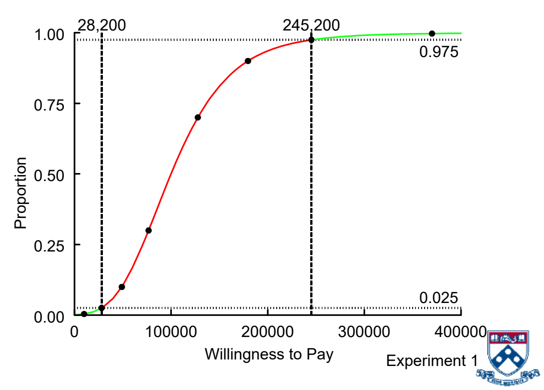
% Acceptable, W = 245,200



Constructing the Acceptability Curve



Acceptability Curve



“Common” Conclusions from Acceptability Curves

W	What is often said
28,200	“97.5% chance Rx A not good value” (Rx B good value)
76,800	“70% chance Rx A not good value”
100,000	“50% chance either therapy good value”
127,700	“70% chance Rx A good value” (Rx B not good value
245,200	“97.5% chance Rx A good value”

- Common to adopt 1-tailed interpretation of acceptability curve
- Ignores fact that 50% – not 0% – represents no information for distinguishing between therapies

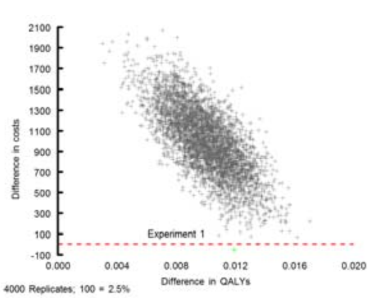
2-tailed Confidence Statements

- Two-tailed confidence statements
  - (For heights > 0.5) Confidence level:

    (2 * Height) - 1
    - e.g., if height of curve is 0.975 for W = 50,000, (2 * .975) -1 = “95% confident that therapy is acceptable / cost-effective”
  - (For heights < 0.5) Confidence level:

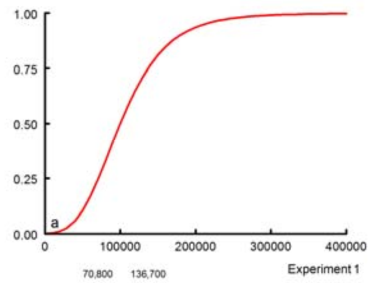
    1-(2*Height)
    - e.g., if height of curve is 0.025 for W = 50,000, “95% confident alternative therapy is acceptable / cost-effective”

Acceptability Curve Additional Information (1)



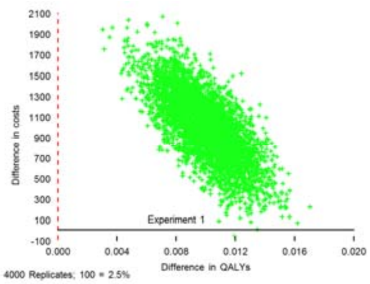
- Counting the points above and below 0 on the x axis tells us proportion acceptable when WTP=0
- What else have we concluded after counting points above and below 0 on the x axis (slide 10)?

Additional Information: 1-Tailed P-Value, ΔC



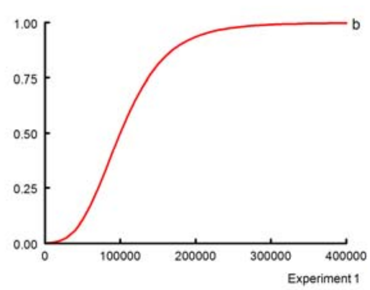
- a: curve height at 0 on x axis = 1-tailed p, ΔC (if curve height $\le$ 0.5)
- 1-curve height at 0 = 1-tailed p, ΔC (if curve height $\ge$ 0.5)

Additional Information (2)



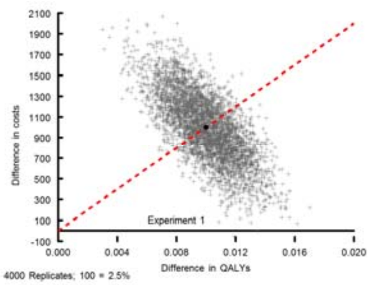
- Counting points to right and left of 0 on Y-axis tells us proportion acceptable when WTP=∞
- What else have we concluded after counting points to right and left of 0 on the Y-axis (slide 8)?

Additional Information: 1-Tailed P-Value, ΔQ



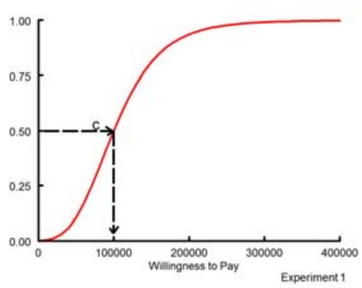
- b: curve height as WTP approaches ∞ = 1-tailed p, ΔQ (if curve height $\le$ 0.5)
- 1-curve height as WTP approaches ∞ = 1-tailed p, ΔQ (if curve height $\ge$ 0.5)

Additional Information (3)



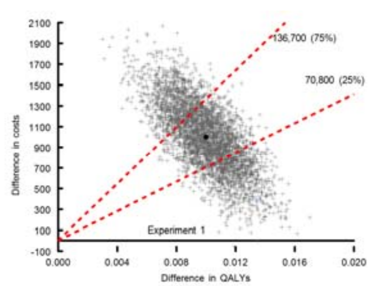
- Parametrically, line through point estimate of ICER defines 50% acceptable
- Nonparametrically, line through point estimate of ICER approximately defines 50% acceptable

Additional Information, Point Estimate of ICER



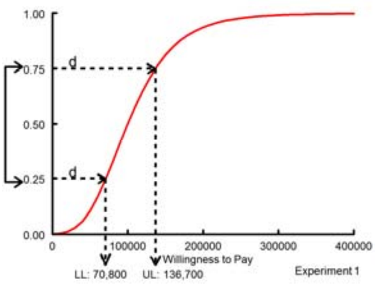
- c: for parametric curve, value of W where curve height = 0.5 represents point estimate of ICER
- For nonparametric curve, value of W where curve height = 0.5 approximates point estimate of ICER

Additional Information (4)



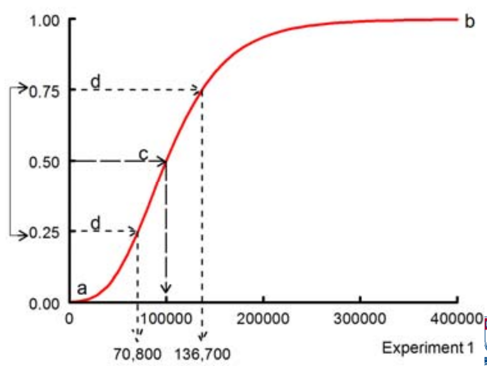
- Values of W for lines through the origin that cut-off 25% (70,800) and 75% (136,700) of distribution represent confidence limits for 50% CI

Additional Information, CL for ICER

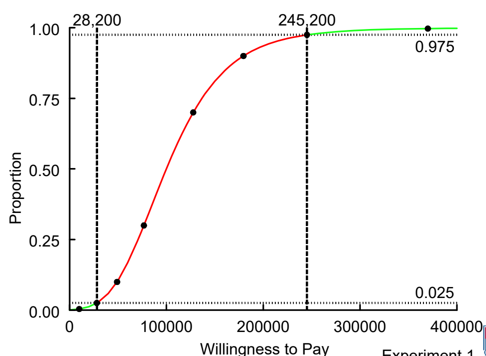


- d: values of W where curve heights = 0.25 (70,800) and 0.75 (136,700) represent 50% confidence limits
- More generally, values of W where curve heights = α/2 and 1-α/2 represent α% confidence limits

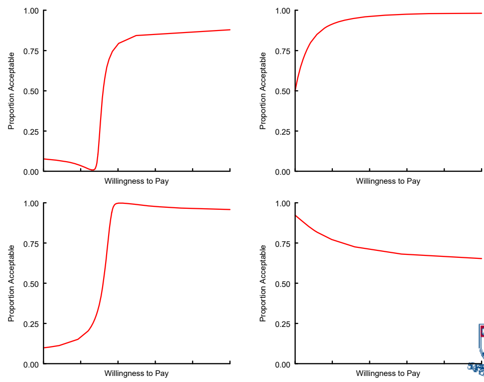
Additional Information in Acceptability Curve



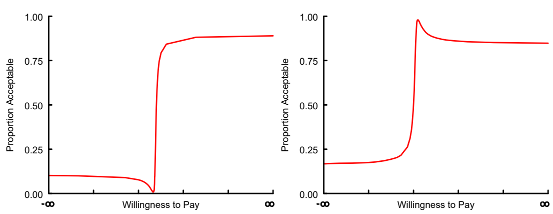
Common to Believe Curves Always Look Like This



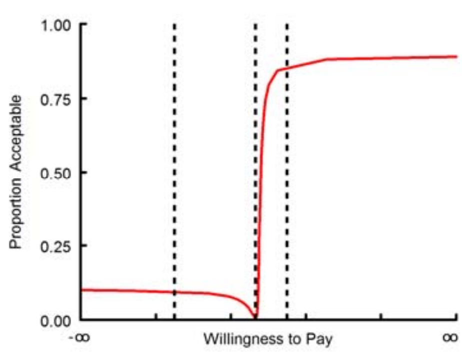
Observable Acceptability Curves for WTP ≥ 0

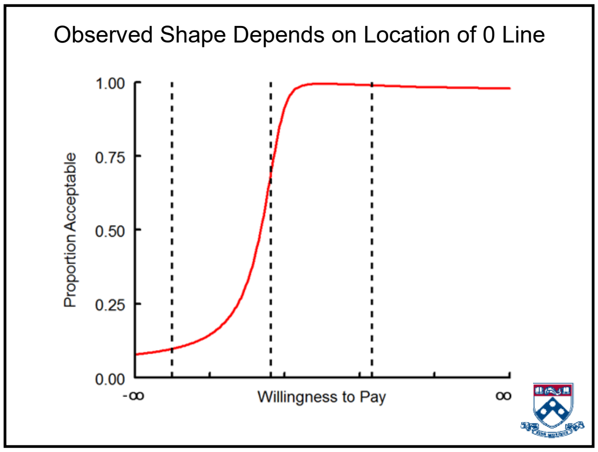


Two Basic Acceptability Curve Patterns



Observed Shape Depends on Location of 0 Line




Observed Shape Depends on Location of 0 Line
1.00
0.75
0.50
0.25
0.00
Proportion Acceptable
-∞
Willingness to Pay
∞