Sampling Uncertainty and Patient-Level Cost-Effectiveness Analysis (Part 2)

Henry A. Glick, Ph.D.

HCMG 901

November 1, 2016

Concerns with CI for ICER

- If every experiment was pattern 1, probably wouldn't have seen development of net monetary benefit and acceptability curves
- But experiments can occur in which CI for ICER have "odd properties" that most people at least initially find counter-intuitive
  - CI can be undefined
    - Referred to as Pattern 3
  - On real number line, either PE > LL > UL or LL > UL > PE
    - Referred to as Pattern 2

Second Example:

Replicates in all 4 quadrants

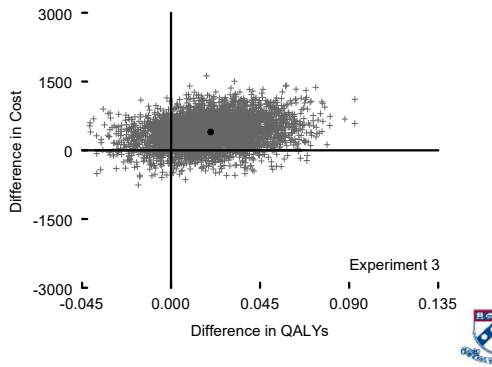
Naïve ordering DOESN'T work

Smart ordering EXTREMELY UNLIKELY TO / DOESN'T work

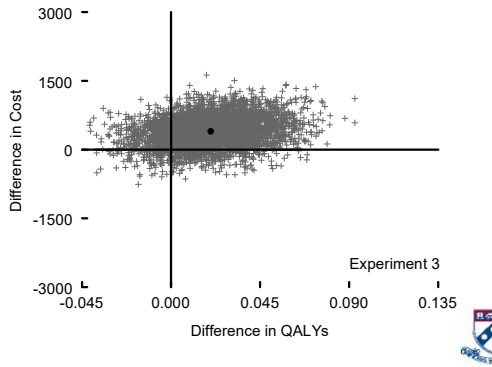
Consider confidence intervals for following experiment:

ΔC=400; SEC=325; ΔQ=.02; SEQ=.02; ρ=0.25; DOF=498

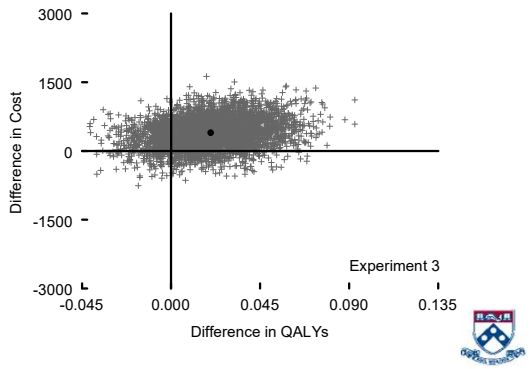
Experiment 3



Order From Lowest to Highest??

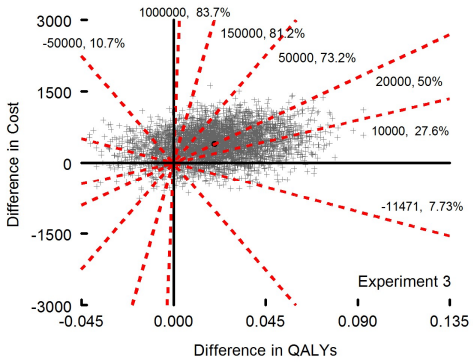


Identify Lines Through Origin That Exclude 2.5%?

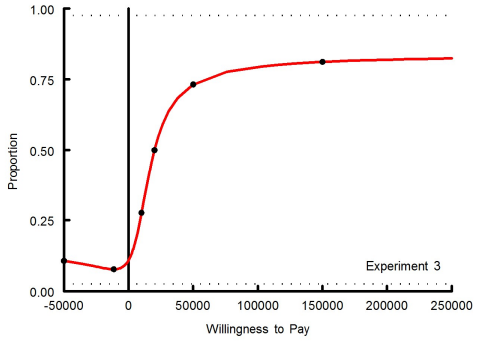


Implications for Acceptability curve?

Can Still Draw Curve

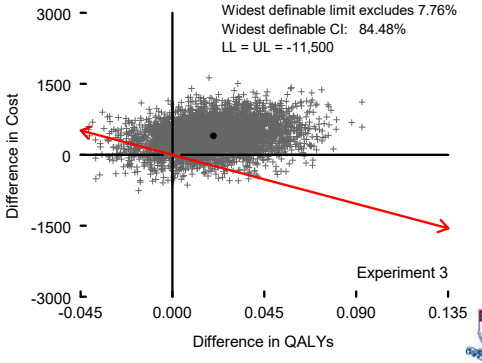


But No Value of W Permits 95% Confidence



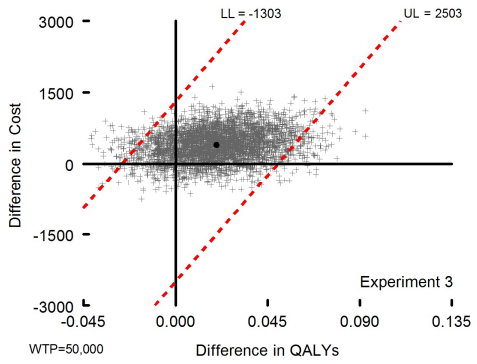
Implications for CI for ICER?

Can Still Define Some CI (eg., 84.48%), But Not 95%



Implications for CI for NMB for Particular W?

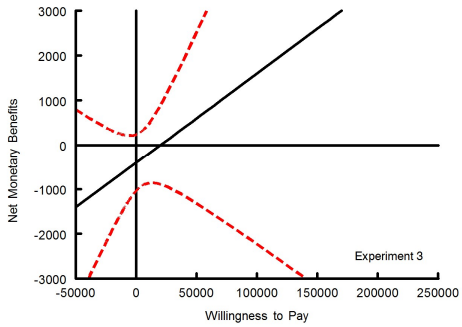
95 CI for NMB for Particular W Always Defined



Implications for NMB Graph?

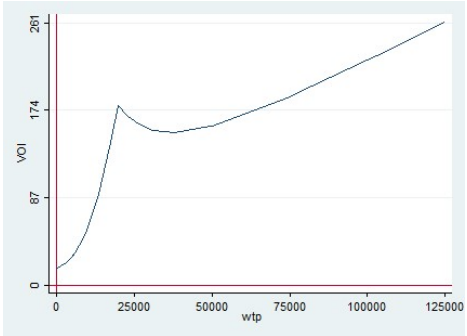
Can Still Draw, But….

No Value of W Permits 95% Confidence

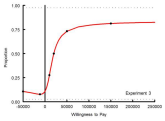


- 0 always falls within 95% confidence interval

Value of Information, Experiment 3

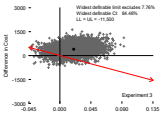


Review of Results for Experiment 3



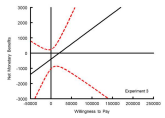
Acceptability curve

Acceptability curve never intersects 0.025 and 0.975 on X-axis



Confidence interval for ICER

95% CI for ICER cannot be defined



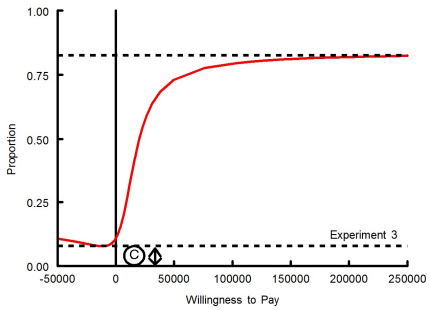
Confidence frontier for NMB

CI never intersect decision threshold (0 NMB / X-axis)

Pattern 3 Findings

- Refer to findings like those in experiment 3 as pattern 3 findings
- 1 of 2 patterns that occur only when difference in effect is not significant
  - P>0.5 for cost necessary but not sufficient condition
- Know we are observing a pattern 3 finding when:
  - Acceptability curve never intersects horizontal lines drawn at either 0.025 and 0.975 on Y axis
  - Confidence interval for the ICER is undefined
  - Neither NMB confidence limit curve intersects decision threshold (0 NMB / X axis)

Region of Acceptability Related to Pattern 3



- For this curve, pattern 3 findings > 84.48% CI

Pattern 3 Findings (2)

Not confident value of two therapies differs

-∞ ⟵ Willingness to Pay ⟶ ∞

Third Example:

Some replicates on both sides of Y-axis, but primarily in 2 or 3 quadrants

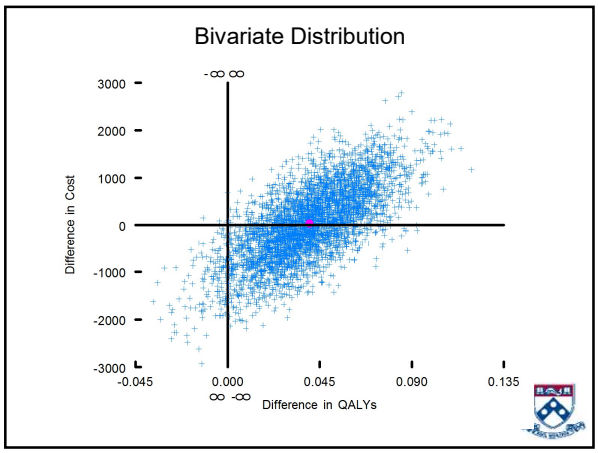
Naïve ordering doesn't work,

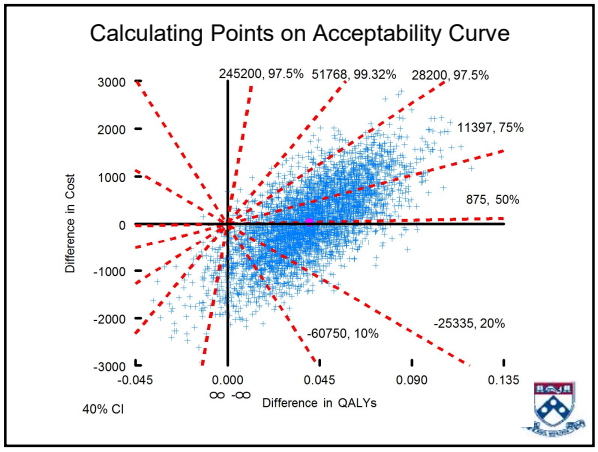
but smart ordering generally does

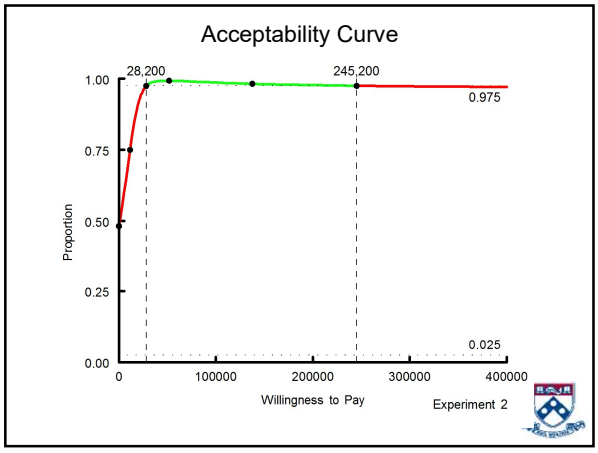
Consider a third experiment that doesn't have either pattern 1 or pattern 3 findings

ΔC=35; SEC=777.06; ΔQ=.04; SEQ=.0224; ρ=0.70625; DOF=498

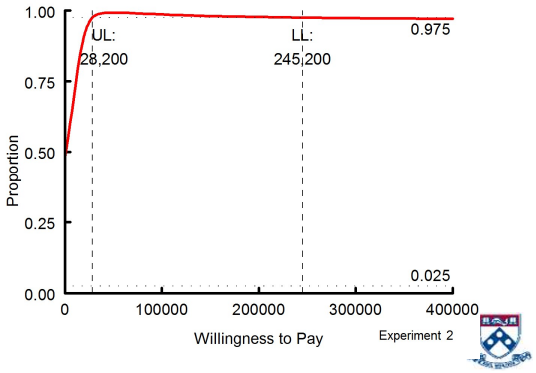
P value for cost, 0.96
P value for QALYs, 0.07
(NEITHER SIGNIFICANTLY DIFFERENT)


Bivariate Distribution
3000
2000
1000
0
-1000
-2000
-3000
Difference in Cost
-∞ ∞
∞ -∞
-0.045
0.000
0.045
0.090
0.135
Difference in QALYs


Calculating Points on Acceptability Curve
245200, 97.5%
51768, 99.32%
28200, 97.5%
11397, 75%
875, 50%
-25335, 20%
-60750, 10%
3000
2000
1000
0
-1000
-2000
-3000
Difference in Cost
-0.045
0.000
0.045
0.090
0.135
Difference in QALYs
40% CI


Acceptability Curve
28,200
245,200
0.975
0.025
1.00
0.75
0.50
0.25
0.00
Proportion
0
100000
200000
300000
400000
Willingness to Pay
Experiment 2

Additional Information, Acceptability Curve

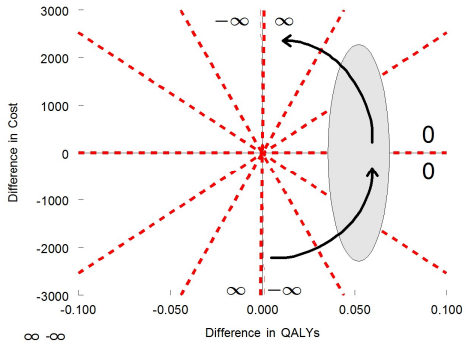


Neither ΔC nor ΔQ significant, but can be 95% confident of value for W between 28,200 and 245,200

For all other values of W can't be 95% confident

CI for ICER When Some Replicates Fall on Each Side of Y Axis?

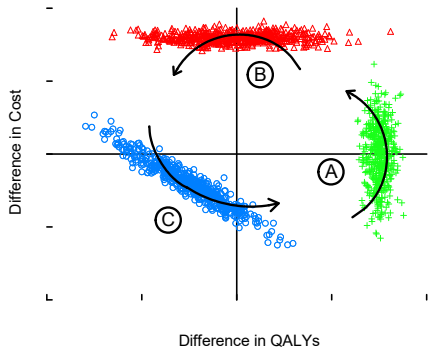
Naïvete Works if Density on One Side of X-Axis



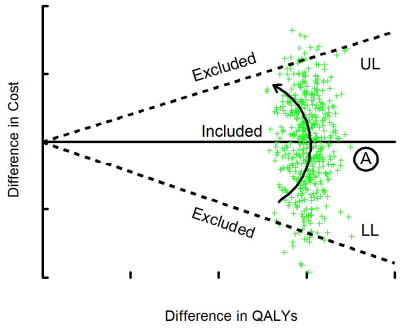
Pluses and Minuses of Ordering for CI for ICER

- But conditions when it fails well defined (e.g., Q: p>.05)
- CI for CER technically **NOT** an "order statistic"
  - Instead defined by lines through origin of CE plane that each exclude α/2% of joint distribution
- **Independent of whether lower limit is a larger or smaller number than upper limit, on CE plane, interval stretches counter-clockwise from lower (clockwise) limit to upper (counter-clockwise) limit**

"Naïve" or "Smart" Ordering Can Work

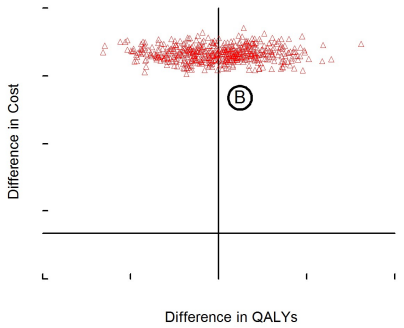


All Replicates in Lower and Upper Right Quadrants

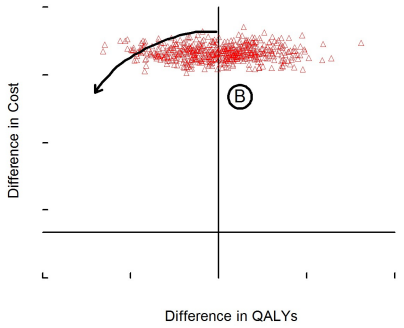


- Naïve ordering (smallest to largest ratio) works

Replicates in Upper Right and Left Quadrants

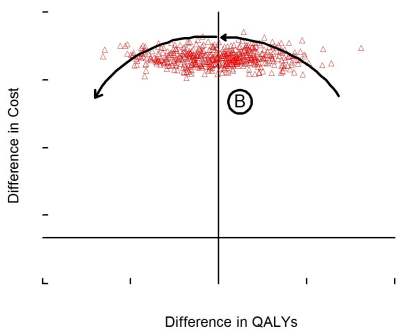


Where Does Naïve Ordering Begin?



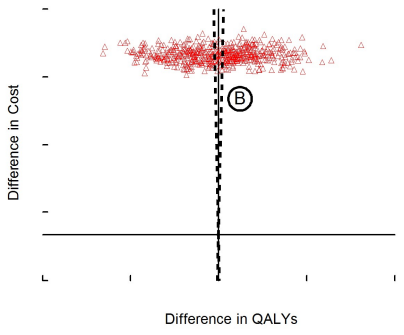
- With smallest negative ratios, left side of Y axis

How Does Naïve Ordering Continue?

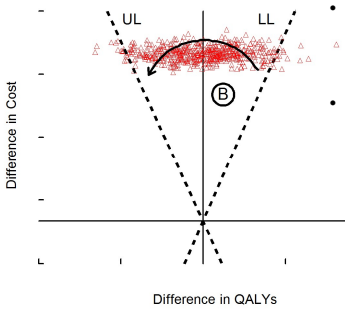


- With smallest positive ratios, right side of Y axis

Cutting off Naively Ordered Smallest and Largest Valued Ratios Omits Wrong Regions



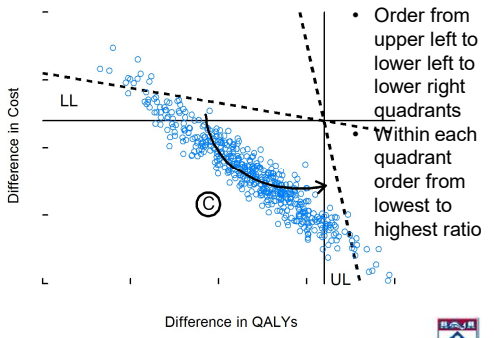
Smart Ordering: Replicates in Upper Right and Left Quadrants



- Order from upper right to upper left quadrants
- Within each quadrant order from lowest to highest ratio

- Smart – but not naïve – ordering works

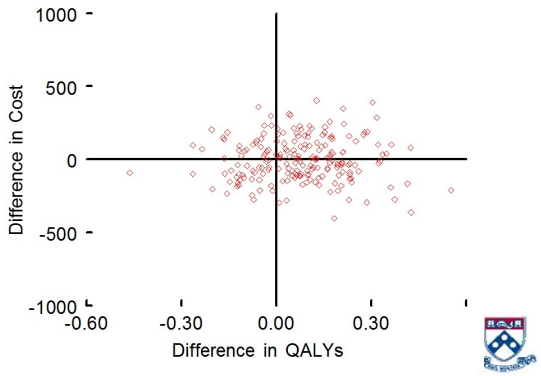
Replicates in 3 Quadrants



- Order from upper left to lower left to lower right quadrants
- Within each quadrant order from lowest to highest ratio

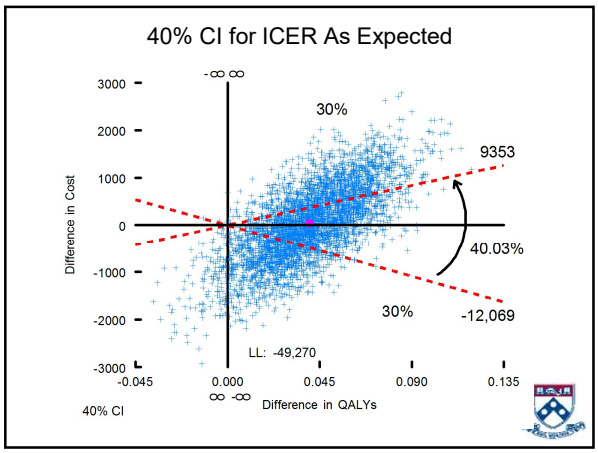
- Smart – but not naïve – ordering also works

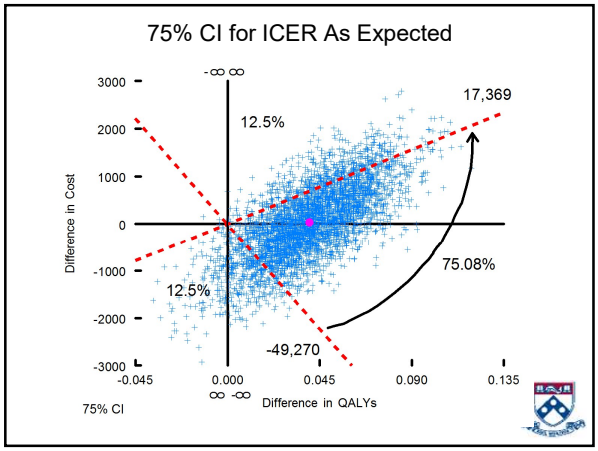
Will Smart Ordering Work?

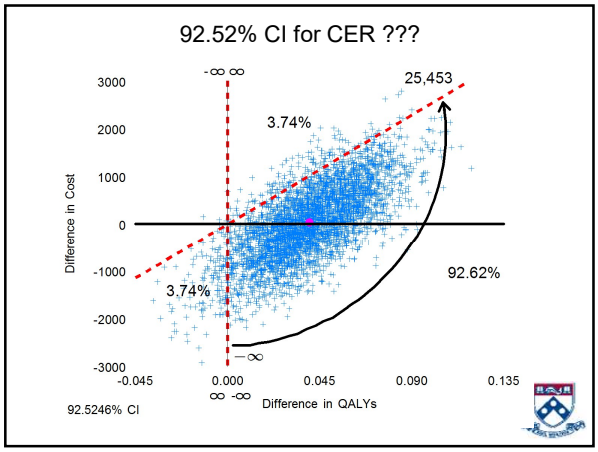


Back to CI for CER for experiment #2

ΔC=35; SEC=777.06; ΔQ=.04; SEQ=.0224;
ρ=0.70625; DOF=498


40% CI for ICER As Expected
30%
9353
40.03%
30%
-12,069
LL: -49,270
Difference in Cost
Difference in QALYs
40% CI


75% CI for ICER As Expected
17,369
12.5%
75.08%
12.5%
-49,270
Difference in Cost
Difference in QALYs
75% CI


92.52% CI for CER ???
25,453
3.74%
92.62%
3.74%
Difference in Cost
Difference in QALYs
92.5246% CI

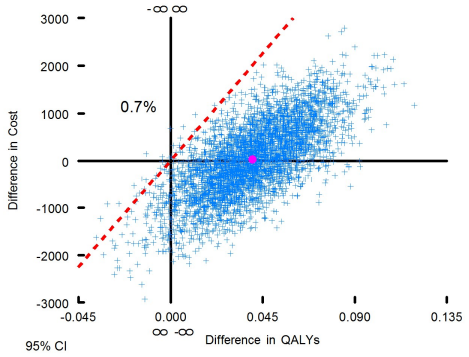
92.52% lower limit equals -∞*

Is 92.52% interval widest that can be defined?

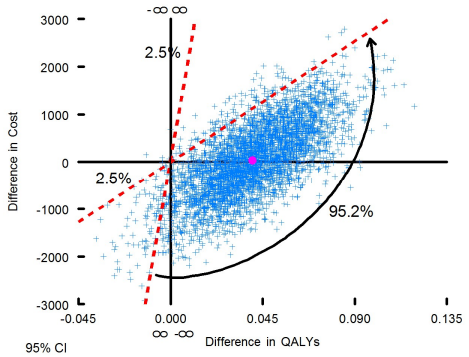
i.e., can we find lines through origin that omit less than 3.74%

(* Technically, lower and upper parametric limits equal +/-∞)

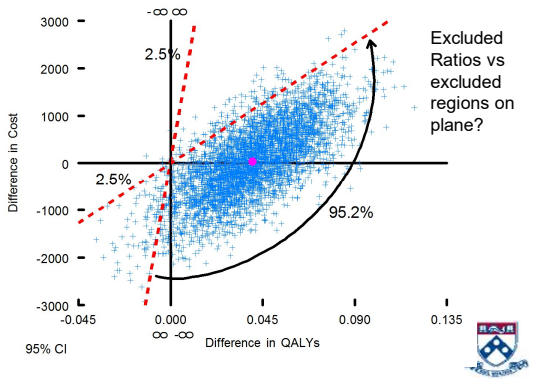
Yes, Lines Omitting < 3.74% Can Be Identified



Thus, Wider Intervals Can be Defined! e.g., 95% CI

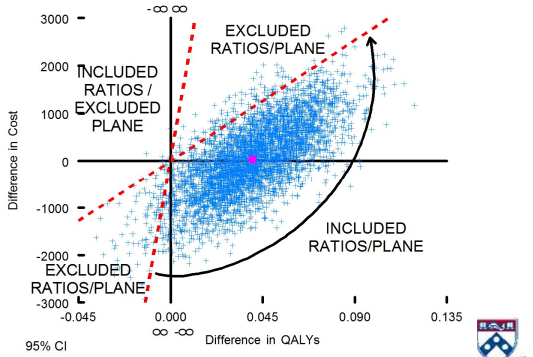


What's Included and What's Excluded?

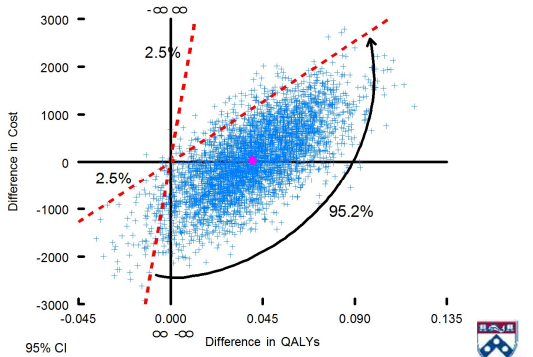


Excluded Ratios vs excluded regions on plane?

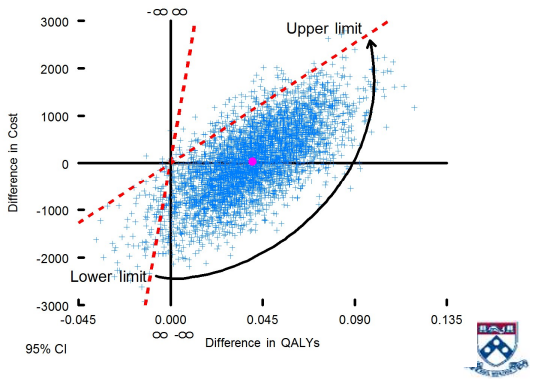
What's Included and What's Excluded?



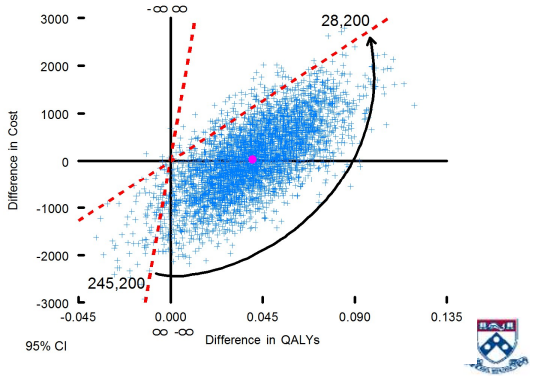
Which is Lower and Which is Upper Limit?



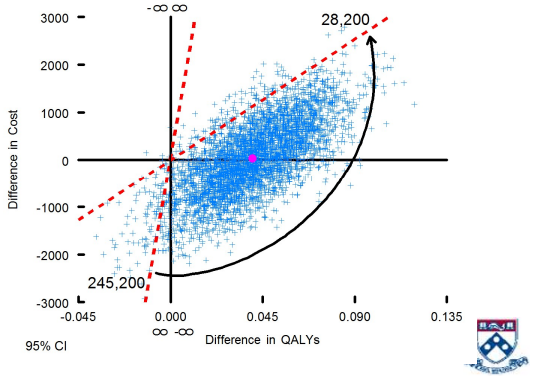
Lower and Upper Limits



95% CI for CER



What Values of WTP Fall Inside Interval?



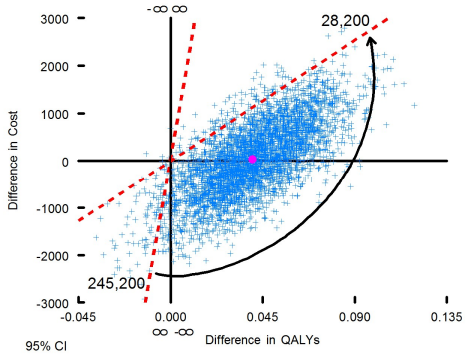
-∞ to 0 (lower right quadrant)

0 to 28,200 (part of upper right quadrant that falls below / to right of upper limit)

245,200 to ∞ (part of lower left quadrant that falls below / to right of lower limit

Interval: -∞ to 28,200 and 245,200 to ∞

What Values of WTP Fall Outside Interval?



28,200 to 245,200 fall outside interval

For what values of WTP can we be 95% confident of value?

If W = 50,000, which therapy is good value?

Confidences Statements for CI for CER

- Confident of value if:
  - P1: LL < UL < W (confident of good value)
  - P1: W < LL < UL (confident of bad value)
  - P2: UL < W < LL (confident of good value if PE<W; confident of bad value if PE>W)
- Not confident of value if:
  - P1: LL < W < UL
  - P2/P3: CI is undefined
  - P2: W < UL < LL
  - P2: UL < LL < W

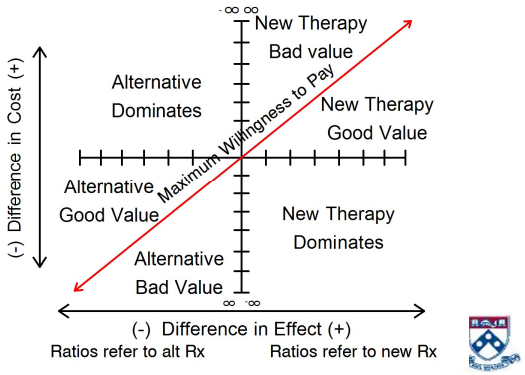
Counter-Intuitive Relationships for CI for ICER

- When more than α/2% of replicates fall on both sides of Y-axis, yet CI is defined:
  - Lower limit (e.g., 245,200) is a larger number than upper limit (e.g., 28,200)
  - ICER point estimate is either a smaller number (e.g., 875 (35/.04)) than both limits or a larger number than both limits
  - Values of WTP included in interval range from -∞ to upper limit and from lower limit to ∞
    - e.g., -∞ to 28,200 and 245,200 to ∞
  - Values of WTP that are excluded from interval range from (smaller) upper limit to (larger) lower limit
  - Confident of value if WTP $\geq$ upper limit and $\leq$ lower limit

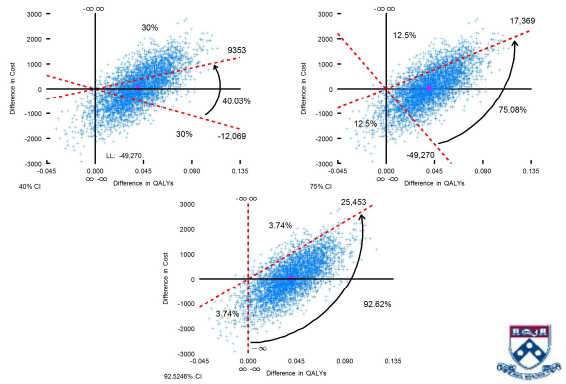
SOURCE OF COUNTER-INTUITIVE RELATIONSHIPS?

On real number line, -∞ and ∞ as far apart as can be

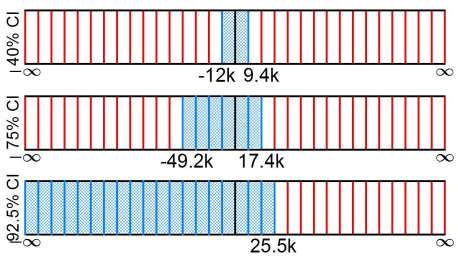
But Infinitesimally Different On Cost-Effectiveness Plane



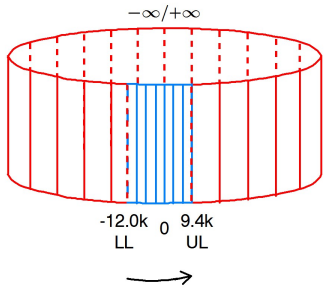
40%, 75% and 92.5% CI on Cost-Effectiveness Plane



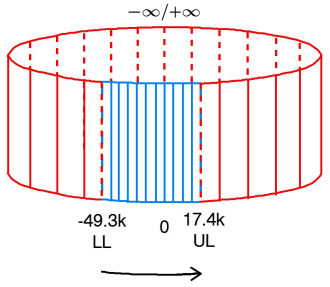
What Happens on Real Number Line



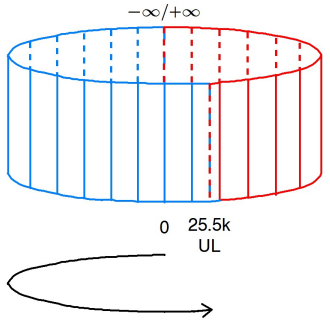
What Happens if We Tape Ends of Real Number Line Together to Form a Ring (40% CI)



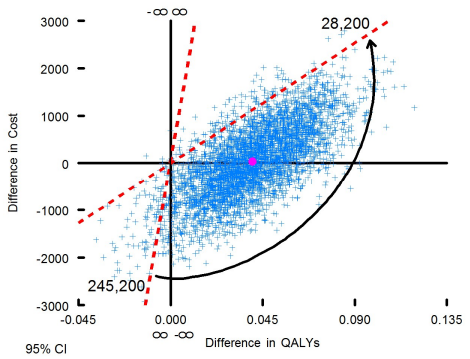
What Happens if We Tape Ends of Real Number Line Together to Form a Ring (75% CI)



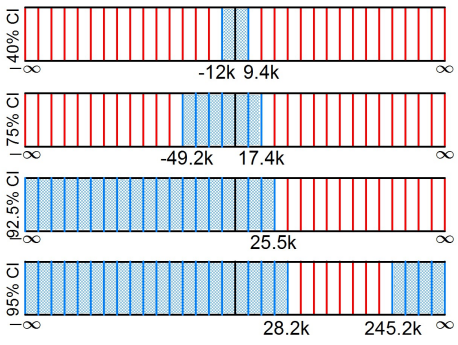
What Happens if We Tape Ends of Real Number Line Together to Form a Ring (92.52% CI)



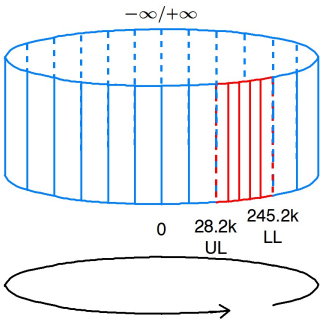
What Happens on Plane When 95% CI Defined?



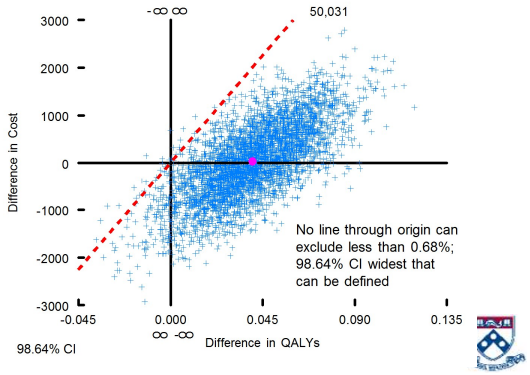
What Happens on Real Number Line?



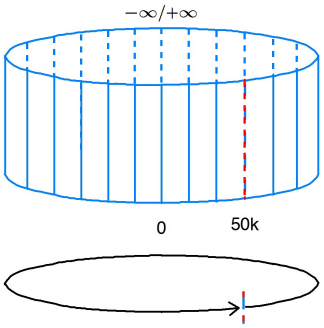
What Happens if We Tape Ends of Real Number Line Together to Form a Ring (95% CI)



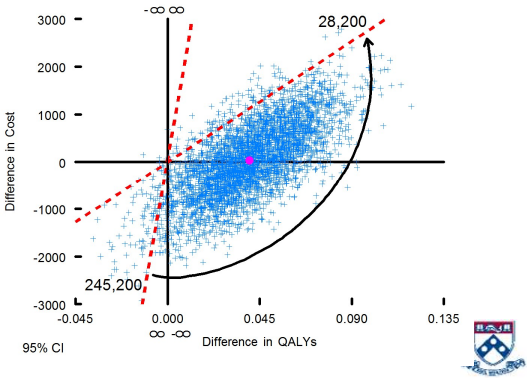
Widest Definable Interval



What Happens if We Tape Ends of Real Number Line Together to Form a Ring (98.6% CI)



95% CI for CER Recap



When Lower Limit is “Larger” than Upper Limit

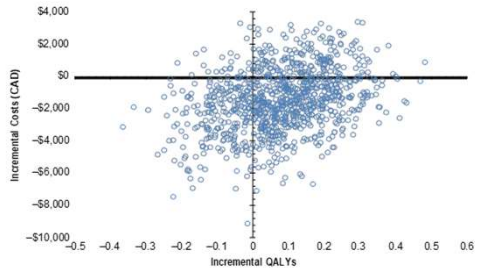
- One of limits indicates that one of therapies may be delivering more health at increased or decreased cost
- Other limit indicates that alternative therapy may be delivering more health at increased or decreased cost
- Q is not statistically significant at α level represented by interval
- Interval thus includes y axis

When Lower Limit is “Larger” than Upper Limit (2)

- Point estimate is either larger than both limits or is smaller than both limits, but does what we expect for one of limits
  - If point estimate and lower limit are on same side of Y axis, point estimate is larger than lower limit (which is larger than upper limit)
  - If point estimate and upper limit are on same side of Y axis, point estimate is smaller than upper limit (which is smaller than lower limit)

Example of Mistakes from Published Literature

Cost-Effectiveness Plane

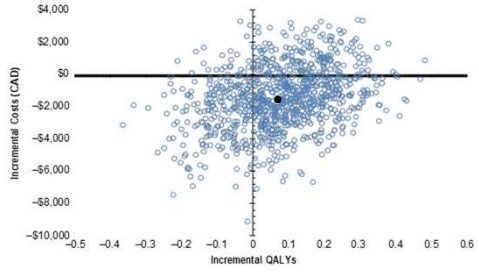


Brown ST, et al. Cost-effectiveness of insulin glargine versus sitagliptin in insulin-naïve patients w/ T2DM. Clin Therapuetics.2014; 36: 1576-87

Brown Table 5

Outcome	Mean	95% CI
Incremental Cost	-1418	-1540 to -1295
Incremental QALYs	0.074	0.066 to 0.082
ICER	-19511	-23,815 to 2044

Authors: ΔC, -1418, 95%CI -1540 to -1295



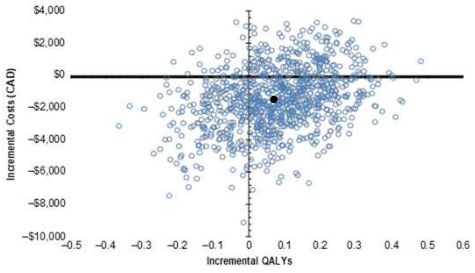
- Elsewhere in paper, authors' report 79% of distribution below X-axis

Too much density on both sides of X-axis
to conclude ΔC significantly
differs from 0

(If 21% above X-Axis, p=0.58)

95% CI cannot equal -1540 to -1295

Authors: ΔQ, 0.074, 95%CI 0.066 to 0.082??



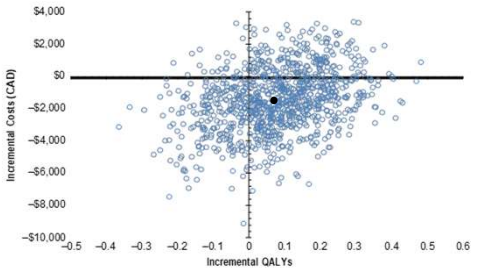
- At least 21% of distribution is to left of Y-axis

Too much density on both sides of Y-axis
to conclude ΔQ significantly
differs from 0

(If 21%+ to left of Y-Axis, p>0.58)

95% CI cannot equal 0.066 to 0.082

Authors: ΔC / ΔQ, -19155, 95%CI -23,815 to $2044



If authors' are correct that costs significantly reduced AND QALYs significantly increased, CI should indicate dominance (e.g., -23,815 to -2044)

Based on scatter plot, cannot identify line through origin that excludes 2.5%

No 95% CI can be defined!

When p>0.05 for ΔQ, lower limit of CI for CER can never be smaller number than upper limit

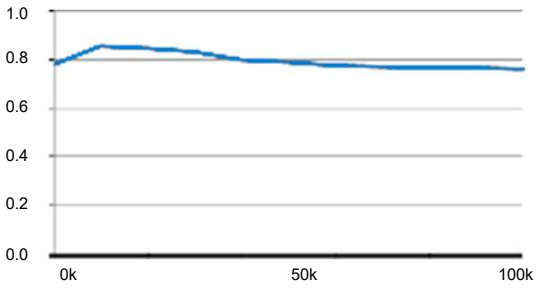
Conclusion

Something very wrong with either Brown et al.'s data plotted on CE plane or with Brown's reported statistics (Probable)

Probably mistakenly divided SE by $N^{½}$

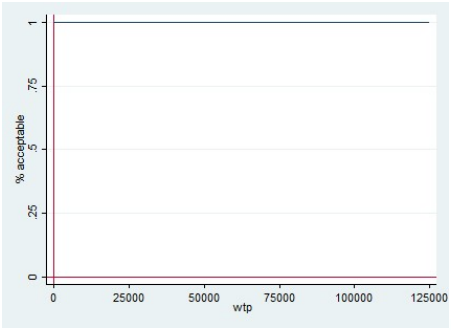
CE Plane does not confirm any statistical conclusions reported in their Table V

Acceptability Curve



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

If Means and SE Were as Suggested…



cd=-1818; sec=62.5; qd=0.074; seq=0.004; (rho=0.1??)

Unique Features of CI for ICER

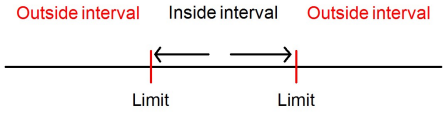
- Role of ∞
  - For OR and RR, widest imaginable limits equal:

    1/∞ to ∞
  - For difference, widest imaginable limits equal:

    -∞ to ∞
  - Do -∞ and +∞ bound the widest CI for an ICER?

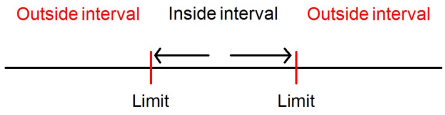
Unique Features (2)

- What's inside and outside the interval?
  - For differences – e.g., NMB, OR, and RR, – what's inside interval ALWAYS falls somewhere in middle of real number line
  - What's outside interval always falls on left and right sides of real number line

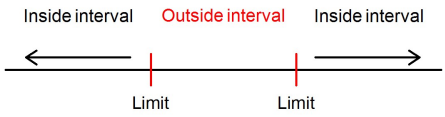


2) What's Inside and Outside Interval?

- For CI for CER, what's inside interval CAN fall somewhere in middle of real number line



- But can also fall on left and right sides of real number line



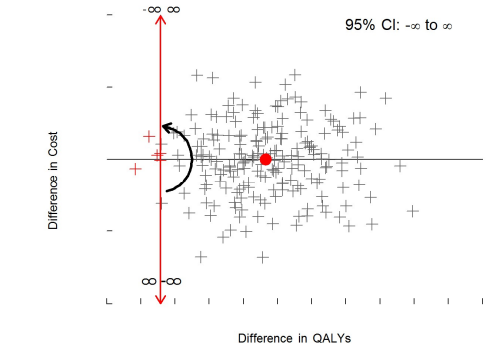
Unique Features (3)

- Reserved numbers for each Rx
  - For differences – e.g., NMB, OR, RR, and acceptability curve – CI has separate ranges of numbers reserved for when one therapy is larger/more effective/more acceptable than alternative versus when it isn't
    - Difference >0, larger than alternative; <0 smaller than alternative
    - OR,RR <1, more effective than alternative; >1, less effective (or vice versa)
    - % acceptable > 0.5 greater likelihood of being good value; <0.5 smaller likelihood of being good value

3) Reserved Numbers

- Numbers reserved for each therapy
  - Not true for CI for ICER
    - When ΔQ>0, CI can include all values between -∞ and ∞

ΔQ>0; 95% CI, -∞ to ∞

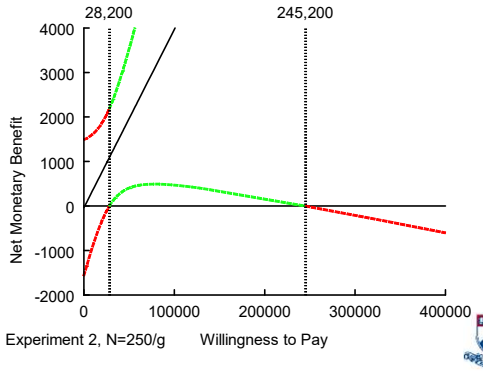


cd=0; se_c=2500; qd=.19569964; se_q=0.1; ρ=0

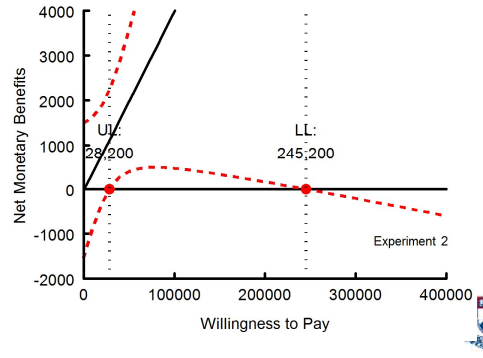
??? Misperceptions About Value When W= ∞ ???

- If Rx deemed good value when WTP = 50k, must it also be good value when WTP=100k? When WTP approaches ∞?
- When designing a study, if power is 50% for WTP=50k and 80% for WTP=100k must power be approaching 100% as WTP approaches ∞?

NMB Graph: 95% Confident for 28,200<WTP<245,200



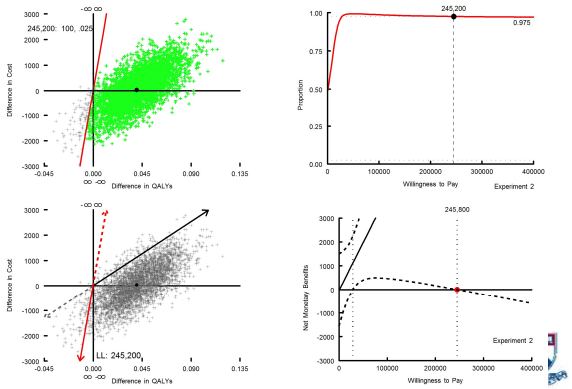
Additional Information, NMB Graph



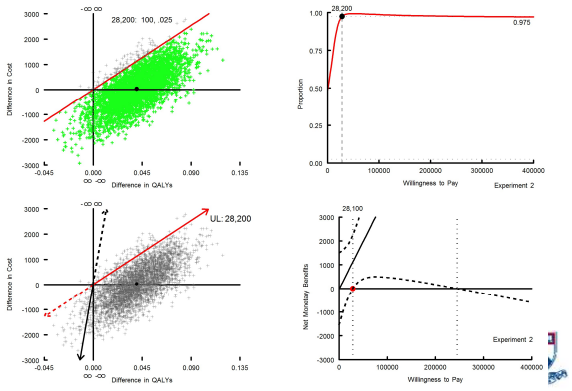
Confidences Statements for CI for NMB

- If both confidence limits negative, 95% confident therapy is bad value
  - In this experiment, does not occur
- If both confidence limits positive, 95% confident therapy is good value
  - i.e., for values of WTP $\geq$ 28,200 and $\leq$ 245,200
- If one confidence limit positive and one negative, cannot be 95% confident value of 2 therapies differs
  - i.e., for values of WTP < 28,200 and > 245,200

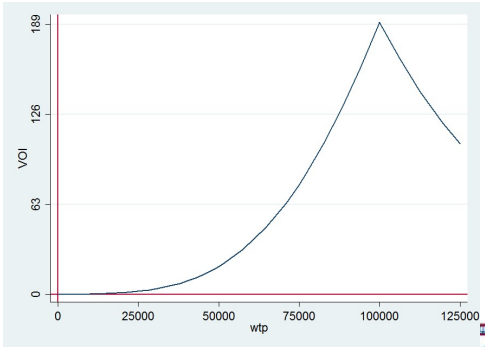
CI for ICER, CI for NMB, Acc Curve All Use Same Lines



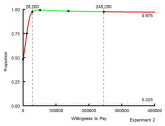
CI for ICER, CI for NMB, Acc Curve All Use Same Lines



Value of Information, Experiment 2

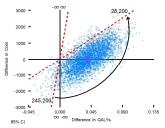


Review of Results for Experiment 2



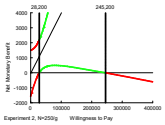
Acceptability curve

Acceptability curve intersects line at 0.975 at 28,200 and 245,200



Confidence interval for ICER

ICER CI: (-∞ to 28,200 and 245,200 to ∞)



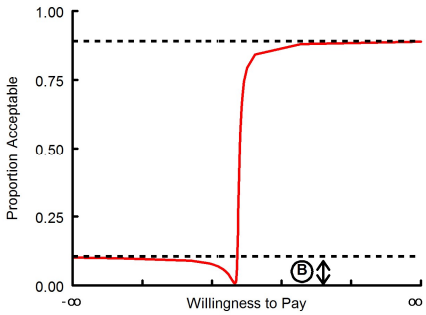
Confidence frontier for NMB

Lower limit intersects decision threshold (0 NMB / X-axis) at 28,200 and 245,200

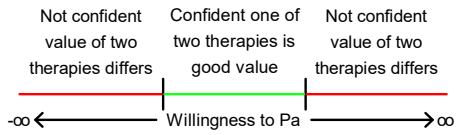
Pattern 2 Findings

- Refer to findings like these as pattern 2 findings
- 1 of 2 patterns that occur only when difference in effect is not significant
- Know we are observing a pattern 2 finding when:
  - Confidence interval for ICER includes Y axis (i.e., LL > UL > PE OR PE > LL > UL)
  - One NMB confidence limit curve intersects decision threshold (0) twice; other limit never intersects decision threshold
  - Acceptability curve intersects a horizontal line drawn at either 0.025 and 0.975 on Y axis twice and never intersects other line

Region of Acceptability Related to Pattern 2



3 Ranges of WTP for Pattern 2 Findings



- In cases where some of boundaries between regions occur at negative values of willingness to pay, may not always observe all 3 regions on acceptability curve or NMB plot

Conclusions (1)

- For any given W, an experiment **ALWAYS** supports one of three conclusions:
  - Confident one therapy good value compared to alternative
  - Confident alternative therapy good value compared to first
  - Cannot be confident that two therapies differ in economic value

Conclusions (2)

- If goal is to identify which of 3 statements holds for a given W, confidence intervals for cost-effectiveness ratios, confidence intervals for NMB, and acceptability curves **ALWAYS** provide same answer
  - e.g., if W included within CI for CER, then:
    - CI for NMB that is calculated by use of W will include 0, and
    - Fraction of distribution that is acceptable at W will fall between horizontal lines that define decision threshold (e.g., between 0.025 and 0.975)

Conclusions (3)

- Confidence intervals for cost-effectiveness ratios provide concise information (i.e., 0, 1, or 2 numbers) that allows determination – based on a particular W – of confidence about a therapy's value
- Acceptability curves provide added advantage of allowing decision makers to assess alternate levels of confidence if such alternate levels are of interest