Sampling Uncertainty and Patient-Level Cost-Effectiveness Analysis

Society for Medical Decision Making Short Course #AM7

October 20, 2012

Henry Glick

www.uphs.upenn.edu/dgimhsr

Good Value for the Cost

- Common goal of an economic analysis is to identify when we can be confident that one therapy is good value compared to another
- One threat to such confidence arises because the economic result observed in an experiment may not truly reflect the result in the population
  - Single sample drawn from a 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

Outline

- Describe methods for identifying when we can and cannot be confident about a therapy's value for the cost
  - Point estimates
  - Confidence intervals
  - Decision threshold
- Goal is to demonstrate the quantification and interpretation of sampling certainty by use of CI for CER, CI for NMB, and acceptability curves
- Don't focus on the technical aspects of estimation

Sampling Uncertainty and Clinical Outcomes

- We can be confident that a therapy is clinically effective when its confidence interval excludes our decision threshold; we can't be confident when its interval includes our decision threshold
- For odds ratios / relative risks. Decision threshold = ???
  - OR = 0.30; 95% CI, 0.15 to 0.63
  - OR = 0.30; 95% CI, 0.09 to 1.02
- For risk differences or changes in blood pressure or cholesterol. Decision threshold = ???
  - Risk difference = 30%; 95% CI, 18% to 42%
  - Risk difference = 30%, 95% CI, -4% to 64%

Implications

- If the confidence interval **includes** the decision threshold, we **CANNOT** be confident that the alternatives differ from one another
- If the confidence interval **excludes** the decision threshold, we **CAN** be confident that the alternatives differ from one another
- It doesn't matter what else is included or excluded from the interval

Sampling Uncertainty and Economic Outcomes

- Confidence statements about economic outcomes are also based on whether or not the confidence interval for the economic outcome includes the decision threshold
- Methods for assessing confidence
  - Confidence intervals for cost-effectiveness ratios
  - Confidence intervals for net monetary benefits
  - Acceptability curve

Confidence Statements

- 95% CI for ICER
  - If the interval excludes W, confident that the two interventions differ in their cost effectiveness
- 95% CI for NMB
  - If the interval excludes 0, confident that the two interventions differ in their NMB
- Acceptability curve
  - If the curve is below α/2 or above 1-(α/2), confident that the two interventions differ in their cost-effectiveness

CI Issues

- What is the threshold, maximum willingness to pay?
  - Differs across jurisdictions
  - Differs within jurisdictions
- Should we be 95% confident?

Conclusions (1)

- For any given willingness to pay, an experiment **ALWAYS** allows us to draw one of three conclusions:
  - We can be confident that one therapy is good value compared to the alternative
  - We can be confident that the alternative therapy is good value compared to the first
  - We cannot be confident that the two therapies differ in their economic value

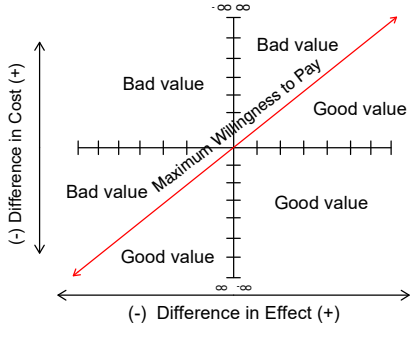
Conclusions (2)

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

Conclusions (3)

- Confidence intervals for cost-effectiveness ratios provide decision makers with concise information (i.e., 0, 1, or 2 numbers) that allows them to determine – based on their own WTP -- if they can be confident about a therapy's value
- Acceptability curves provide the added advantage of allowing decision makers to assess alternate levels of confidence if such alternate levels are of interest

Cost-Effectiveness Plane



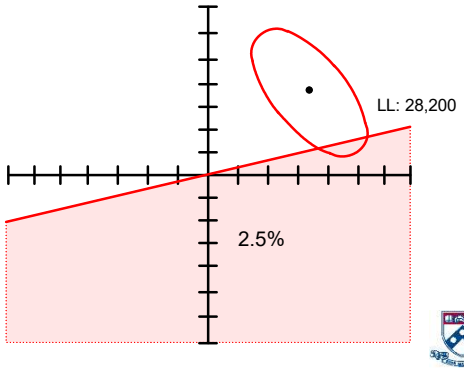
Example

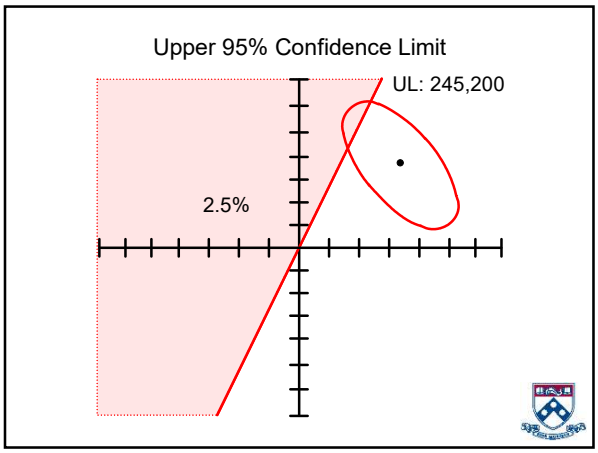
- Two therapies, A and B
- Δcost = 1000 (SE: 325, p=0.002)
- ΔQALYs = 0.01 (SE: 0.001925, p=0.000)
- A is significantly more costly and significantly more effective
  - CER = 1000 / 0.01 = 100,000 / QALY saved
- 250 participants in each arm of the trial
- Correlation between cost and effect is -0.71

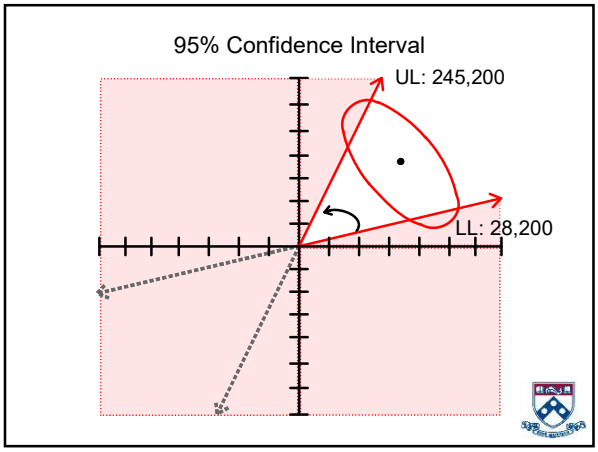
Confidence Intervals for CER

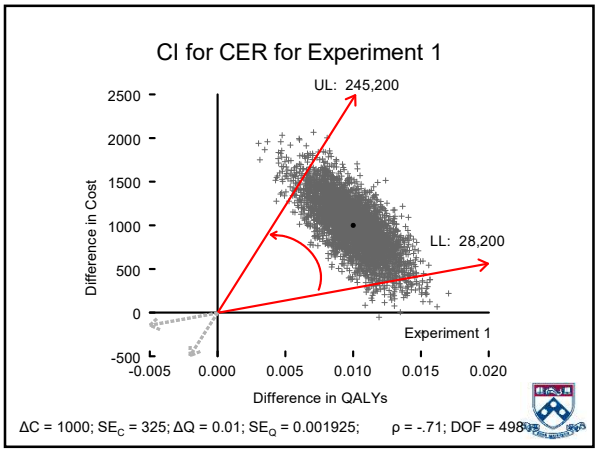
- Common wisdom: order ratios from lowest to highest; ICERS in the replicates representing the 2.5th and 97.5th percentiles represent the lower and upper limits
  - "Order statistic"
- CI for CER technically not an order statistic (although in many cases is equivalent to one)
- Lines through the origin that exclude α/2 of the joint distribution of the difference in cost and effect
- Independent of whether the lower limit is a larger or smaller number than the upper limit, on cost-effectiveness plane, interval stretches counter-clockwise from lower (clockwise) limit to the upper (counter-clockwise) limit

Lower 95% Confidence Limit




Upper 95% Confidence Limit
UL: 245,200
2.5%


95% Confidence Interval
UL: 245,200
LL: 28,200


CI for CER for Experiment 1
UL: 245,200
LL: 28,200
Difference in Cost
2500
2000
1500
1000
500
0
-500
-0.005
0.000
0.005
0.010
0.015
0.020
Experiment 1
Difference in QALYs
ΔC = 1000; SE_C = 325; ΔQ = 0.01; SE_Q = 0.001925; ρ = -.71; DOF = 498

Confidences Statements for CI for CER

- Can be confident of value when W not included in the confidence interval
- When lower limit is a smaller number than upper limit
  - Interval ranges between lower and upper limit
    - 28,200 to 245,200
  - Confident of value if WTP is either smaller than lower limit or greater than upper limit
    - Confident of bad value if WTP < 28,200
      - Because at least 97.5% of the samples have ratios greater than 28,200
    - Confident of good value if WTP > 245,200
      - Because at least 97.5% of the samples have ratios less than 245,200

NMB Recap

$$NMB = (W^*\Delta Q) - \Delta C$$

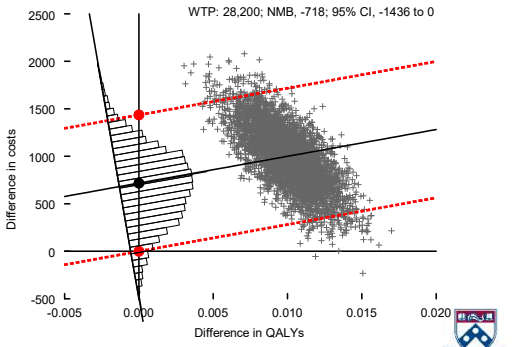
- For a WTP of 50,000, NMB for experiment 1:

  (50,000 * .01) -1000 = -500
- The study result is a difference in means of net benefits, not a ratio of means, and is always defined (i.e., no odd statistical properties like the ratio) and continuous
- Unlike the cost-effectiveness ratio, the standard error of net benefits is always defined
- Given that not all decision making bodies have an agreed upon maximum willingness to pay, we routinely estimate net benefit over the range of policy relevant values of willingness to pay

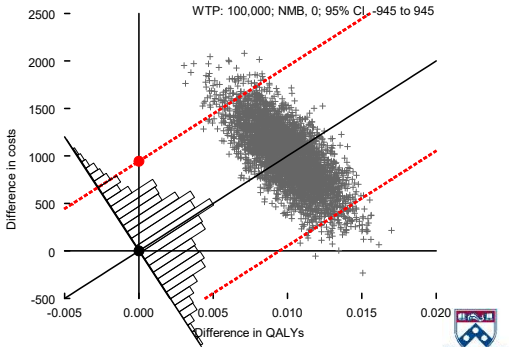
Net Benefit Graphically

- Defined on the cost effectiveness plane using a family of lines
- Each line represents a single value of NMB and equals −intercept (because when $\Delta Q=0$, $W\Delta Q$ drops out of the equation
- Slope of all lines equal to W
- 95% CI for NMB defined by identifying the 2 NMB lines that each omit 2.5% of the distribution

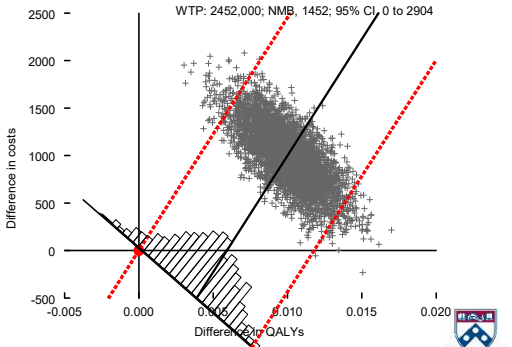
Constructing CI for NMB for Experiment 1, WTP=28.2K



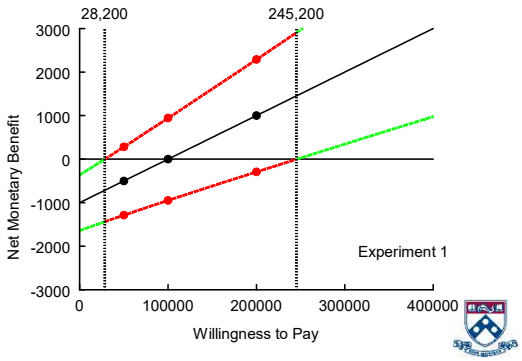
Constructing CI for NMB for Experiment 1, WTP=100K



Constructing CI for NMB for Experiment 1, WTP=245.2K



CI for NMB for Multiple WTP

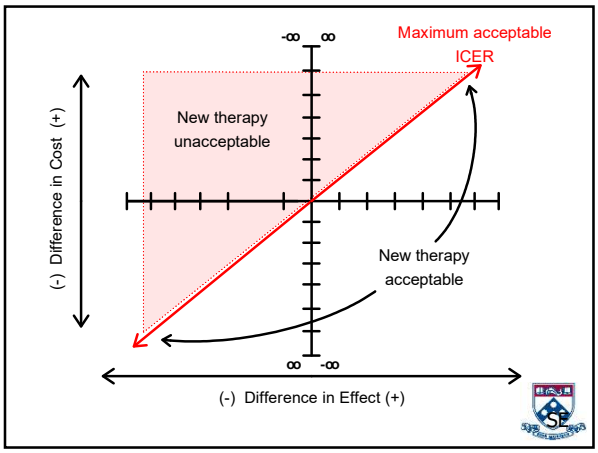


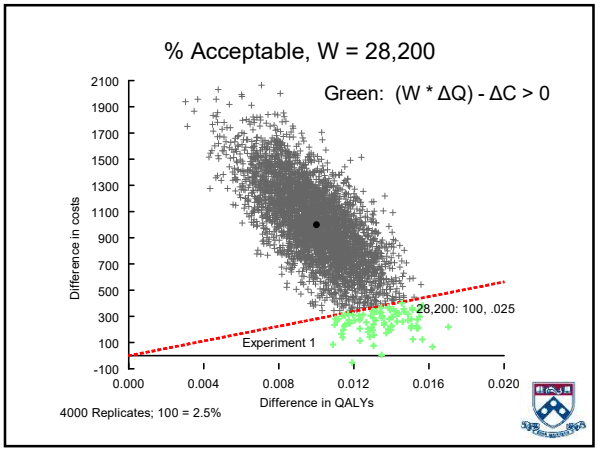
Confidences Statements for CI for NMB

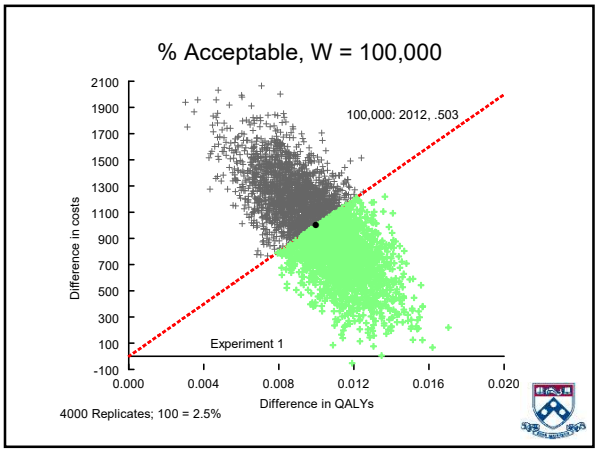
- If both confidence limits are negative, 95% confident the therapy is bad value
  - i.e., for values of WTP $\leq$ 28,200
- If both confidence limits are positive, 95% confident the therapy is good value
  - i.e., for values of WTP $\geq$ 245,200
- If one confidence limit is positive and one is negative, cannot be 95% confident the value of the 2 therapies differ
  - i.e., for values of WTP > 28,200 and < 245,200

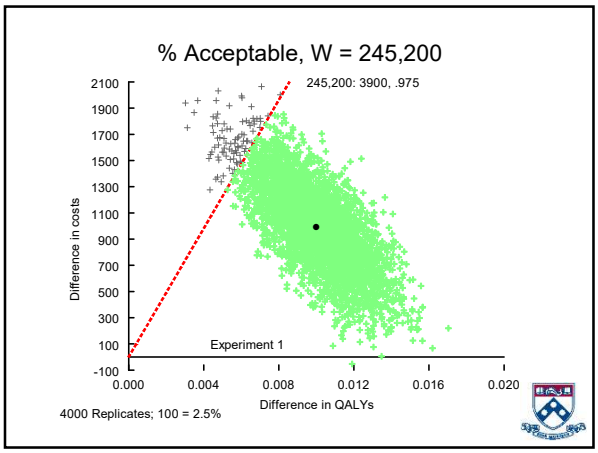
Acceptability Curve

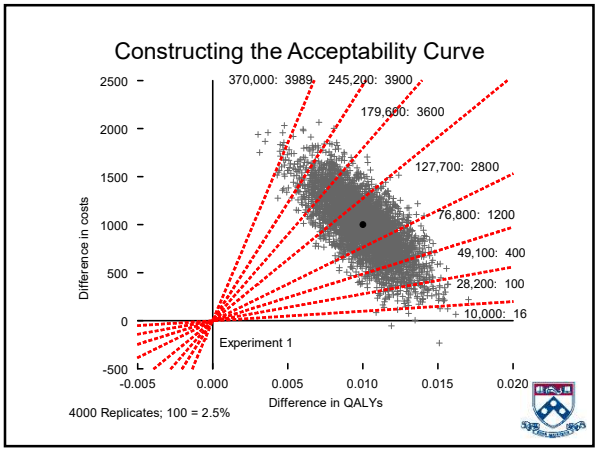
- Acceptability criterion defined on the cost-effectiveness plane as a line passing through the origin with a slope equal to WTP
- Proportion of the distribution of the difference in cost and effect that falls below and to the right of this line is "acceptable" (i.e., has positive NMB)
  - Proportion acceptable for one therapy = 1-proportion acceptable for alternative therapy
- Proportion that is above and to the left of this line is "unacceptable“
  - Proportion unacceptable for one therapy = 1-proportion unacceptable for alternative therapy

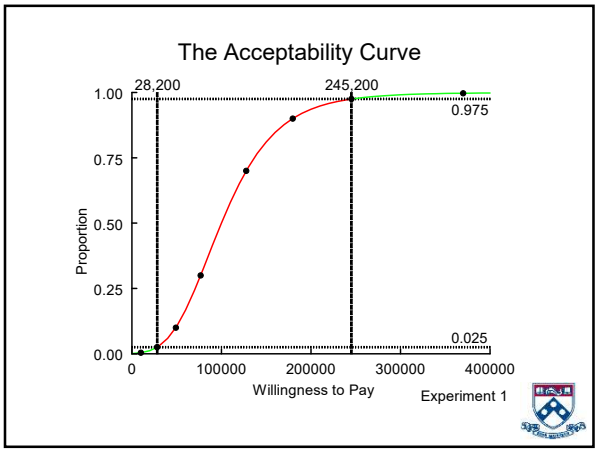

Maximum acceptable ICER
New therapy unacceptable
New therapy acceptable
(-) Difference in Cost (+)
(-) Difference in Effect (+)
-∞
∞
∞
-∞


% Acceptable, W = 28,200
Green: (W * ΔQ) - ΔC > 0
Difference in costs
Difference in QALYs
28,200: 100, .025
Experiment 1
4000 Replicates; 100 = 2.5%

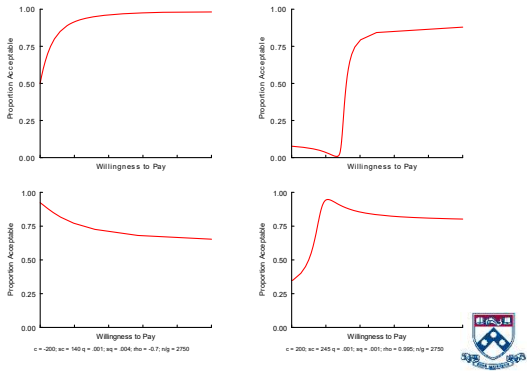

% Acceptable, W = 100,000
100,000: 2012, .503
Difference in costs
Difference in QALYs
Experiment 1
4000 Replicates; 100 = 2.5%


% Acceptable, W = 245,200
245,200: 3900, .975
Difference in costs
2100
1900
1700
1500
1300
1100
900
700
500
300
100
-100
Experiment 1
0.000
0.004
0.008
0.012
0.016
0.020
Difference in QALYs
4000 Replicates; 100 = 2.5%

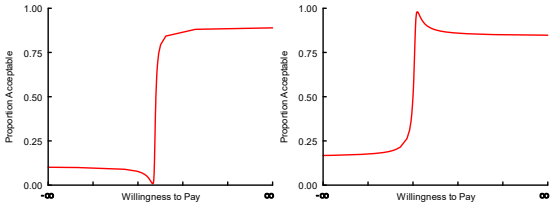

Constructing the Acceptability Curve
370,000: 3989
245,200: 3900
179,600: 3600
127,700: 2800
76,800: 1200
49,100: 400
28,200: 100
10,000: 16
Difference in costs
2500
2000
1500
1000
500
0
-500
Experiment 1
-0.005
0.000
0.005
0.010
0.015
0.020
Difference in QALYs
4000 Replicates; 100 = 2.5%


The Acceptability Curve
28,200
245,200
0.975
0.025
Proportion
1.00
0.75
0.50
0.25
0.00
0
100000
200000
300000
400000
Willingness to Pay
Experiment 1

Observable Acceptabilty Curves for WTP ≥ 0



Two Basic Patterns for Acceptabilty Curves



2-tailed Confidence Statements

- (For heights > 0.5) Confidence level:

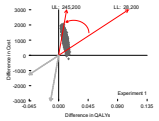
  1 - (2*(1-Height))

  – e.g., if the curve has a height of 0.975 for W = 50,000, 1 - (2*(1-.975)) = "95 confident that the therapy is good value / cost-effective / acceptable"
- (For heights < 0.5) Confidence level:

  1-(2*Height)

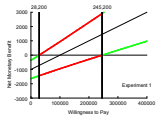
  – e.g., if the curve has a height of 0.025 for W = 50,000, "95% confident the alternative therapy is acceptable / cost-effective"

Review of Results for Experiment 1



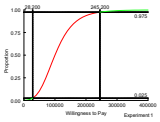
Confidence interval for CER

CER CI: (28,200 to 245,200)



Confidence frontier for NMB

CI intersect decision threshold (0) at 28,200 to 245,200



Acceptability curve

Acceptability curve intersects 0.025 and 0.975 at 28,200 and 245,200

"Pattern 1" Findings

- We refer to findings like those in experiment 1 as pattern 1 findings
- They occur when the difference in effect is significant
- We know we are observing a pattern 1 finding when:
  - The confidence interval for the cost-effectiveness ratio excludes the Y axis (i.e., LL < PE < UL)
  - Both NMB confidence limits curves intersect the decision threshold (0) once
  - The acceptability curve intersects horizontal lines drawn at both 0.025 and 0.975 on the Y axis

3 Ranges of WTP for Pattern 1 Findings

One can be confident the more effective therapy is not good value	One cannot be confident the two therapies differ from one another	One can be confident the more effective therapy is good value

-∞ ← Willingness to Pay → ∞

* In cases where some of the boundaries between the regions occur at negative willingnesses to pay, we may not always observe all 3 regions on an acceptability curve or NMB plot

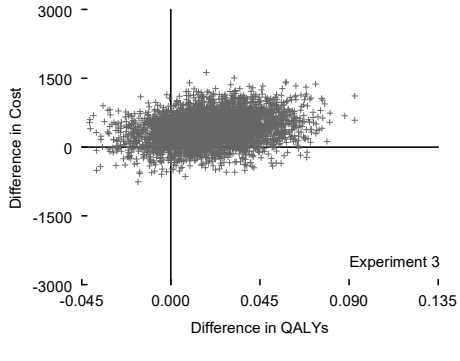
Concerns With CI for ICER

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

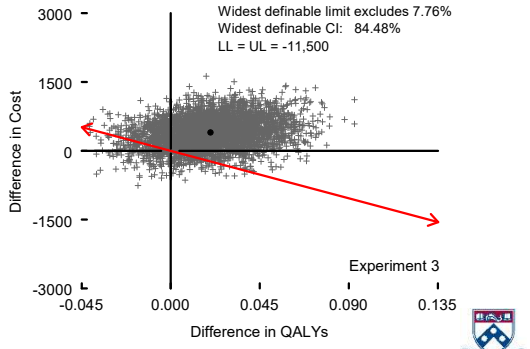
Consider the confidence intervals for the following experiment:

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

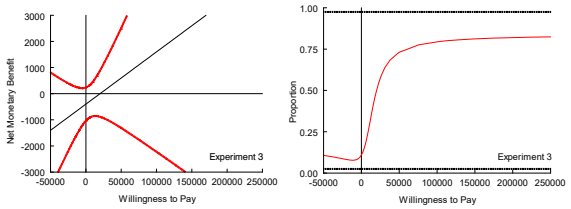
Experiment 3



Widest Definable Interval, Experiment 3



CI for NMB and Acceptability Curve



Pattern 3 Findings

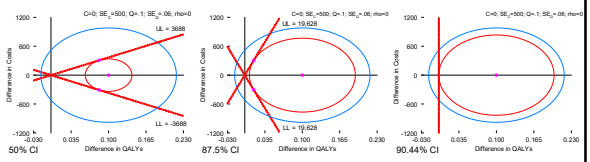
- We refer to findings like those in experiment 3 as pattern 3 findings
- They occur only when the difference in effect is not significant
- We know are observing a pattern 3 finding when:
  - The confidence interval for the CER is undefined
  - Neither NMB confidence limit curve intersects the decision threshold (0)
  - The acceptability curve never intersects horizontal lines drawn at either 0.025 and 0.975 on the Y axis

1 Range of WTP for Pattern 3 Findings

We cannot be confident the two therapies differ from one another

-∞ ⟵ Willingness to Pay ⟶ ∞

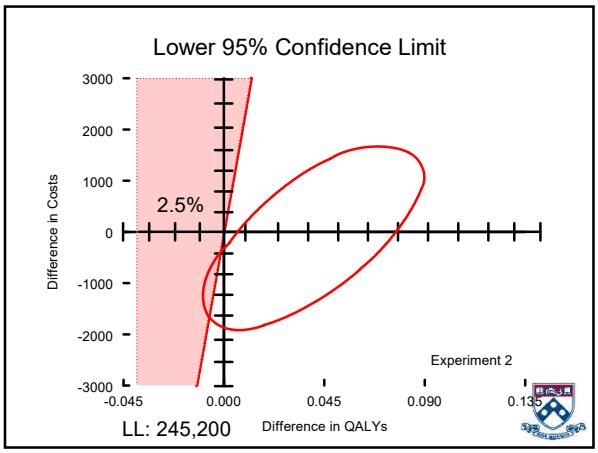
Confidence Intervals When ΔQ is Significant

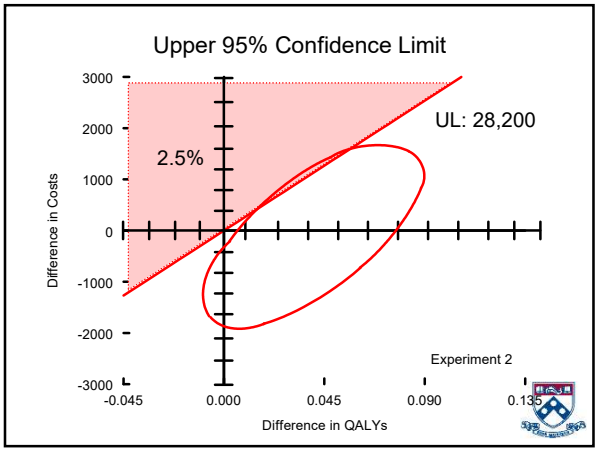


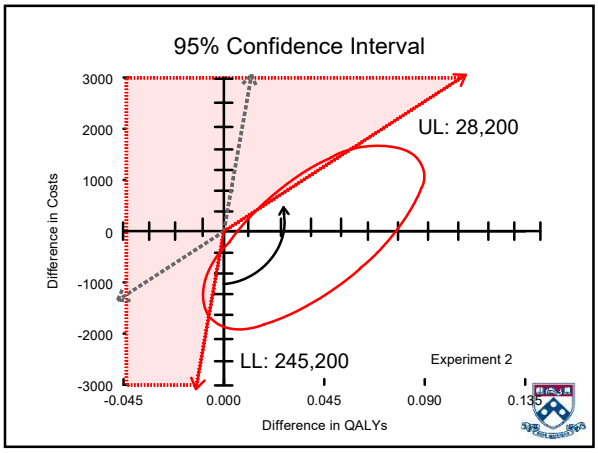
- CI can include all real numbers from -∞ to ∞ when ΔQ is significant
- What convention will we use so we know when an experiment indicates that ΔQ is not significant?

Consider the confidence intervals for the following experiment:

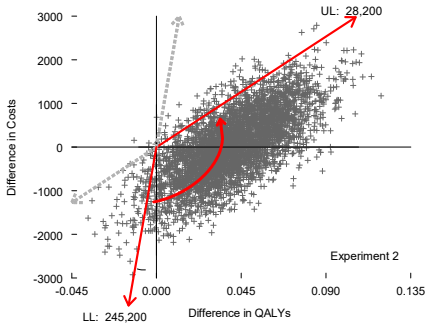
ΔC=35; SEC=777.5; ΔQ=.04; SEQ=.0224; ρ=0.706; DOF=498


Lower 95% Confidence Limit
3000
2000
1000
0
-1000
-2000
-3000
Difference in Costs
2.5%
Experiment 2
-0.045
0.000
0.045
0.090
0.135
Difference in QALYs
LL: 245,200


Upper 95% Confidence Limit
3000
2000
1000
0
-1000
-2000
-3000
Difference in Costs
2.5%
UL: 28,200
Experiment 2
-0.045
0.000
0.045
0.090
0.135
Difference in QALYs

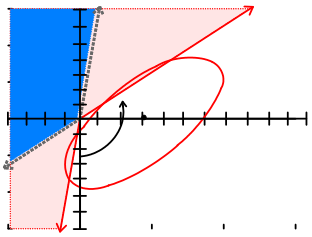

95% Confidence Interval
3000
2000
1000
0
-1000
-2000
-3000
Difference in Costs
UL: 28,200
LL: 245,200
Experiment 2
-0.045
0.000
0.045
0.090
0.135
Difference in QALYs

CI for CER for Experiment 2



$\Delta C = 35$; $SE_C = 777.5$; $\Delta Q = 0.04$; $SE_Q = 0.0224$; $\rho = .706$; DOF = 498

Fraction of Distribution Excluded from Interval



- If each limit excludes α/2 and the intersection of the area excluded by the 2 limits need not be empty, what is the implication for the fraction of distribution that is jointly excluded by the 1-α% CI?

Confidences Statements for CI for CER

- Can be confident of value when W not included in the confidence interval
- When lower limit is a larger number than the upper limit
  - Interval ranges between -∞ and upper limit and between lower limit and ∞
  - Values of W between upper limit and lower limit are excluded from the interval
  - Confident of value if WTP is larger than upper limit and smaller than lower limit

When the Lower Limit is Larger than the Upper Limit

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

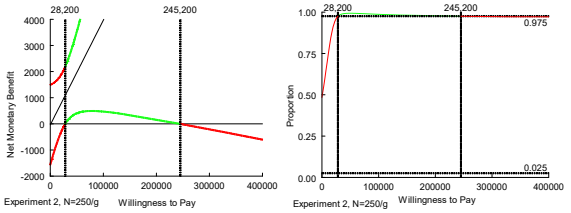
When Lower Limit is "Larger" than Upper Limit (2)

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

Confidence Statements

- We derive the same confidence statements from CI for CER, CI for NMB, and the acceptability curve
- So long as our WTP is between 28,200 and 245,200, we can be 95% confident that the therapy is good value

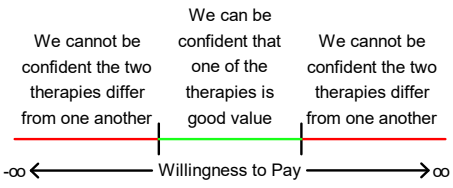
Same Confidence Statements From NMB and Acceptability Curves



Pattern 2 Findings

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

3 Ranges of WTP for Pattern 2 Findings



* In cases where some of the boundaries between the regions occur at negative values of willingness to pay, we will not always observe all 3 regions on an acceptability curve or NMB plot

Calculation

- Programs, datasets, and sample calculations available at:

  www.uphs.upenn.edu/dgimhsr/eeinct_cicer.htm

Conclusions (1)

- For any given willingness to pay, an experiment **ALWAYS** allows us to draw one of three conclusions:
  - We can be confident that one therapy is good value compared to the alternative
  - We can be confident that the alternative therapy is good value compared to the first
  - We cannot be confident that the two therapies differ in their economic value
- If our goal is to identify which of these 3 statements holds for a given willingness to pay, confidence intervals for cost-effectiveness ratios, confidence intervals for NMB, and acceptability curves **ALWAYS** provide the same answer

Conclusions (2)

- Confidence intervals for cost-effectiveness ratios provide decision makers with concise information (i.e., 0, 1, or 2 numbers) that allows them to determine – based on their own WTP -- if they can be confident about a therapy's value
- Acceptability curves provide the added advantage of allowing decision makers to assess alternate levels of confidence if such alternative levels are of interest