The Cost Effectiveness of Cost Effectiveness Analysis in Comparative Effectiveness Research

Henry Bergquist[1], Jalpa A. Doshi[1], Lee Fleisher[1], Henry A. Glick[1], Bruce Kinosian[1], Sean McElligott[1], Eleanor M. Perfetto[2] (Co-PI), Daniel E. Polsky[1], Mark V. Pauly[1] (Co- PI), J. Sanford Schwartz[1] (PI), Richard Willke[2] (Co-PI)

Affiliations: 1 University of Pennsylvania, 2 Pfizer

Funding: Penn-Pfizer Alliance

Motivation

- Comparative effectiveness research (CER) seeks to assess an intervention's incremental clinical benefit
- Economic efficiency requires consideration of real resource costs of alternative interventions. Cost-effectiveness analysis (CEA), as expressed by the incremental cost-effectiveness ratio (ICER), is a well developed method to assess an intervention's incremental costs and benefits
- Decision making based on CER alone can lead to adoption of expensive interventions with only small incremental clinical benefit
- But explicit use of cost information and formal CEA for clinical and policy decision making is contentious in the U.S.

Concerns and Contradictions in US CER / CEA Clinical and Policy Decision Making

- Wilensky: Politically unwise to consider costs explicitly in CER
- Garber: Using CER alone to make decisions like ordering from a menu with no prices
  - May be some circumstances in which costs can be included in CER indirectly
- Current Law: Limits conduct and use CEA in CER (and considerable political pressures not to do so)
  - Selby: "Correct" that PCORI not consider costs
- Federal agencies: USPSTF and Medicare prohibited from consideration of costs and cost-effectiveness in recommendations and policies; ACIP, VA and NIH expert guidelines are not so prohibited

Study Aims

- To address public/political cost-effectiveness "fatigue", Identify predictors of agreement between CER decision making and CEA decision making (i.e., identify circumstances when CEA information provides little incremental value to CER information)
  - Are there systematic characteristics of interventions that we can identify *a priori* that predict when CER decision making is sufficient?
- Empirical analysis: Examine a set of CEA studies to see how frequently and under what circumstances consideration of cost information in conjunction with clinical information lead to the same choice as a decision based on clinical information alone

Study Sample

- Study sample drawn from the Tufts University Center for the Evaluation of Risk in Health CEA Registry

Exclusion Criterion	N
Initial sample	6793
Non-US studies	3718
Studies prior to 1990	87
Missing either ICER or QALYs	1065

TOTAL: 1923 studies

Agreement

- Main outcome: Binary variable representing agreement and disagreement between adoption recommendation from CER and adoption recommendation from CEA
- CER Recommendation: Adopt therapy with the larger point estimate for effectiveness
  - In most formal research, effectiveness measures will be disease-specific clinical outcomes
    - e.g., changes in HbA1c, mm/Hg of blood pressure, or mmol/l of cholesterol
  - In current study, effectiveness measure is QALYs derived from the denominator of the cost/QALY ratio

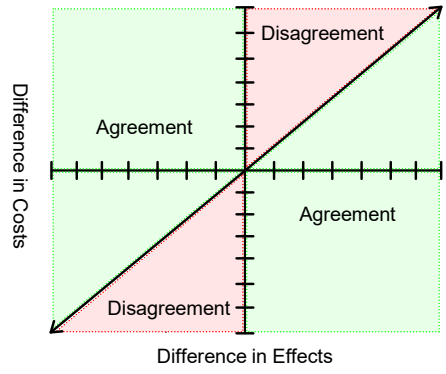
Agreement (2)

- Cost-effectiveness analysis: Compares difference in cost with difference in effect between pairs of therapies

$$ICER = (C_1 + C_0) / (E_1 + E_0)$$

- Ratio generally interpreted as the extra payment per extra unit of effectiveness for more effective therapy
- CEA Recommendation: Adopt therapy that is good value based on the point estimate for the ICER (2010 US$ per QALY) and a pre-specified WTP threshold
  - $100,000 per QALY with $50,000 sensitivity analysis
    - e.g., Adopt more effective therapy if it does more and costs less than alternative or has an ICER<100,000
    - Adopt less effective therapy if more effective therapy has an ICER>100,00

Don't Distinguish Between "Types" of Agreement



Explanatory variables

- Type of intervention (9)
  - Surgical (index), care delivery, device, diagnostic, health education, medical procedure, pharmaceutical, screening, other
- Disease category (12)
  - Cardiovascular (index), infectious, musculoskeletal, maternal health, sensory organ, cancer, environmental, mental health, digestive, respiratory, endocrine, other
- Source of funding
  - Industry vs other

Explanatory variables (2)

- Prevention stage
  - Primary: Methods used to prevent disease or illness
  - Secondary: Methods used to diagnose and treat disease in early stages before causing significant morbidity
  - Tertiary: Methods used to reduce negative impact of disease by restoring function and reducing disease-related complications

Explanatory variables (3)

- "Publicness" of disease
  - Google trends: Relative search volume for all 71 study conditions in the CEA Registry
- Year of study
- Research "intensity"
  - Clinicaltrials.gov: Mapped 93,722 clinical studies by 289 MeSH terms into 12 major disease categories
    - Used the number of studies in each disease category to define inverse probability weights that were used in all models
      - Smaller number of studies, more weight

Logistic Regression Agreement Models

- 3 Logistic regression models
- Model 1: Agreement as a function of
  - Type of intervention
  - Disease category
  - Prevention stage
  - Funding source
  - "Publicness of disease"
- Model 2: Model 1 + year fixed effects
- Model 3: Model 2 + interactions between diseases and Google trends "Publicness"
- Coefficients are odds ratios (OR < 1 CER/CEA more likely to disagree; OR > 1 CER/CEA more likely to agree)

Slide 12

HAG8 Are these all of the variables? What about disease burden and research intensity?
Henry Glick, 6/13/2012

Unadjusted Agreement, Overall and By Type of Intervention

	Agree	~Agree
Overall, N, (%)	1338 (70)	585
Type of intervention (p=0.001)		
Most agreement		
Other	17 (85)	3
Care delivery	44 (83)	9
Device	47 (78)	13
Least agreement		
Pharmaceutical	552 (68)	256
Screening	193 (64)	107
Diagnostic	55 (56)	44

Unadjusted Agreement, Disease Category

	Agree	~Agree
Disease category (p<0.0001)		
Most agreement		
Cardiovascular	293 (76)	95
Sensory organs	44 (75)	15
Infectious & parasitic	203 (74)	71
Least agreement		
Respiratory	40 (59)	28
Musculoskeletal	80 (58)	57
Maternal & perinatal	16 (57)	12

Unadjusted Agreement, Prevention Stage, Funding Source, and Year

	Agree	~Agree
Prevention Stage (p=0.06)		
Primary	244 (67)	120
Secondary	356 (67)	177
Tertiary	738 (72)	288
Funding Source (p<0.0001)		
Industry	295 (79)	77
Other	1043 (67)	508
Year (p=0.81)		
Pre 2000	403	173
Post 1999	935	412

Unadjusted Agreement, Publicness, Disease Burden, and Research Intensity

Variable, Mean (SD)	Agree	~Agree	P-value
"Publicness"	0.34	0.39	0.22
Research Intensity	0.012	0.010	0.01

Odds Ratios from Agreement Logits *

Characteristic	Model 1	**Model 2**	**Model 3**
Type of intervention			
Pharmaceutical	***0.546 §***	***0.534 §***	***0.422 §***
Disease group			
Cancer	0.650 †	***0.622 ‡***	0.693
Musculoskeletal	***0.581 ‡***	***0.539 ‡***	***0.473 §***
Neuropsychiatric	0.791	0.705	***0.576 ‡***
Prevention stage			
Secondary	1.818 †	1.835 †	***2.082 ‡***
Tertiary	1.753 †	***1.872 ‡***	***1.900 ‡***

* >1 = more agreement; <1 = more disagreement; robust standard errors clustered at the article level

† p < 0.1, ‡ p < 0.05, § p < 0.01

$50,000 Sensitivity Analysis

- Pharmaceutical interventions: Decreased likelihood of agreement more strongly significant in all 3 models
- Screening: Significantly less agreement in all 3 models
- Musculoskeletal and cancer remain significant for lower agreement
- Infectious and parasitic diseases: Significantly greater agreement
- Neuropsychiatric diseases and prevention stage generally no longer significant

Increased Disagreement for Pharmaceuticals

- Not sure what is causing, but
  - Reasonable spread of pharmaceutical studies across disease areas
    - Reduces likelihood that result due to pharmaceutical studies being clumped in a few disease areas for which WTP might be substantially higher than the $100k threshold we adopted
  - Pharmaceutical studies tend to concentrate in tertiary prevention
    - Tertiary care has greater agreement, but pharmaceuticals have less agreement in this environment of greater agreement

Limitations

- Are CEA analyses included in our study representative of clinical decisions for which CER analyses will be performed?
  - Publication bias?
- Agreement between CER measured by comparing incremental QALYs and incremental cost per QALY ratios
- Could not account for uncertainty
- Single WTP threshold

Incremental QALYs vs Cost per QALY

- Does the 70% agreement we observed between QALY gains (CER) and cost per QALY ratios (CEA) translate to agreement for other outcomes such as biomarkersl?
  - e.g., Simply knowing drug A reduces cholesterol more than drug B does not imply drug A's resulting increase in QALYs makes it good or bad value
- Depends in part on whether studies in which cost-effectiveness has been reported are a representative sample of studies in which CER will be performed

Uncertainty

- CEA Registry does not report variability of the difference in costs or effects or of the cost-effectiveness ratio
- Addition of variability generally thought to increase agreement
  - Point estimates indicate disagreement, but one or both estimates not significant (no significant difference in effectiveness or CI for CER that includes WTP) and we cannot be confident of disagreement
- But can decrease agreement
  - Point estimates indicate agreement, but nonsignificance of one or both estimates reduces confidence of agreement

Single WTP Threshold

- Given US thresholds generally unknown, difficult to evaluate use of different WTP thresholds for different diseases
- Allowing different diseases to have different thresholds generally thought to increase agreement
  - e.g., if treatments for musculoskeletal or neuropsychiatric diseases or primary prevention have WTP thresholds >100,000
- But also can decrease agreement
  - Do some diseases have lower values of WTP?
  - Interaction with uncertainty?
    - Possible to have less certainty of value as WTP approaches ∞

"Correct" That PCORI Not Consider Costs?

- No evidence that adopting the more effective therapy saves health care $
  - In 72% of the studies in our sample, the more effective therapy was associated with higher costs
  - 28% with lower costs probably overstates the likelihood of savings in health care $
    - A number of studies derived savings from non-health care $ (e.g., work loss)

"Correct" That PCORI Not Consider Costs? (2)

- Rationale: PCORI should "put the emphasis on clinical outcomes" and local public and private decision makers can develop economic evidence
  - Can't be efficient
  - Quality of evidence will be mixed at best
- Does development of clinical but not economic evidence make controlling costs harder rather than easier?
  - "But PCORI reported its the most effective therapy...."
  - Future legislation?: "Insurers must cover the most effective therapy as determined by PCORI"
- Should PCORI collect economic data, but not use it in making its recommendations?
  - Would increase efficiency and allow quality monitoring

Conclusions

- Had hoped to be able to develop measurable criteria that allowed us to confidently avoid some CEA so as to avoid CEA fatigue
- Did find that economics data are more likely to raise questions for CER studies of pharmaceuticals, musculoskeletal conditions, and neuropsychiatric conditions as well as for primary prevention
  - Don't appear to be very strong results or rule out many cost-effectiveness analyses

Conclusions (2)

- Large amount of agreement between CER and CEA when QALYs are the outcome measure (i.e., part of the CEA calculation), but:
  - May not translate to studies that use some other CER metric
  - Unclear our findings imply we can avoid CEA for politically visible therapies which probably cause the greatest fatigue
- Still a noble aim, and more research is needed