The Biggest Mistakes in Economic Evaluation

Jeffrey S. Hoch, PhD

Associate Director, Center for Healthcare Policy and Research
Professor and Chief, Division of Health Policy and Management
Department of Public Health Sciences, University of California, Davis
Adjunct Professor, University of Toronto, Canada
Health Economist, St. Michael’s Hospital, Toronto, Ontario, Canada

Main idea: Anatomy of common mistakes

© Jeffrey Hoch, PhD
Main idea: Anatomy of common mistakes (and why they matter)

Create incorrectly

Main idea: Anatomy of common mistakes (and why they matter)

Use incorrectly

© Jeffrey Hoch, PhD
Cascade failure

Create
Bad evidence

Use
Wrong read on the evidence

Result
Poor advice, bad policy, sub-optimal treatment

Main idea: Anatomy of common mistakes

Types of Mistakes

Doing
Data

Using
Analysis

Results

Rule of 2
Rule of Right
Data Mistakes

Rule of 2

• Must collect, analyze and report cost

• and

effect

Rule of “Right”

• Should consider the “right”
  • Perspective
  • Outcome
  • Alternative

Example: mental health

Mental Health Example, II

Data Mistakes

Rule of 2

- Must collect, analyze and report cost
- Must collect, analyze and report effect

Rule of “Right”

- Should consider the “right”
  - Perspective
  - Outcome
  - Alternative
Screening for Prostate Cancer
A Decision Analytic View

Murray D. Kohn, MB, MSc; John E. Mahoney, MD; Mark H. Eckman, MD; et al

Abstract

Objective. —To determine the clinical and economic effects of screening for prostate cancer with prostate-specific antigen (PSA), transrectal ultrasound (TRUS), and digital rectal examination (DRE).

Design. —Decision analytic cost–utility analysis comparing four screening strategies with a strategy of not screening. We assumed that the cancer detection rate and stage distribution were predicted by each combination of tests and that localized cancer was treated with radical prostatectomy. For each strategy, we calculated life expectancy, quality-adjusted life expectancy (QALE), and cost–utility ratios for unscreened and high-prevalence populations.

Data. —Probabilities and rates for clinical events were gathered from published data. We assessed utilities by the time-trade-off method using urologists, radiation oncologists, and internists as subjects. The Clinical Cost Manager at the New England Medical Center provided cost data.

Results. —In unscreened men between the ages of 50 and 70 years, screening with PSA or TRUS prolonged unadjusted life expectancy but diminished QALE. Screening with DRE alone yielded no reduction in mortality at any age. All programs increased costs. Results were sensitive only to assumptions about the efficacy of treatment. In high-prevalence populations, screening produced a similar pattern: gains in unadjusted life expectancy, losses in QALE, and increased costs.

Conclusions. —Our analysis does not support using PSA, TRUS, or DRE to screen asymptomatic men for prostatic cancer. Screening may result in poorer health outcomes and will increase costs dramatically. Assessment of comorbidity, risk attitude, and valuation of sexual function may identify individuals who will benefit from screening, but selecting high-prevalence populations will not improve the benefit of screening (JAMA. 1994;272:773-780).

“In unselected men between... 50-70 years, screening with PSA... prolonged unadjusted life expectancy but diminished quality-adjusted life expectancy (QALE)
Data Mistakes

**Rule of 2**

- Must collect, analyze and report **cost**
- **and**  **effect**

**Rule of “Right”**

- Should consider the “right”
  - Perspective
  - Outcome
  - Alternative
What about change in year, location or situation?

Data Mistakes

Rule of 2

- Must collect, analyze and report cost
- and
- effect

Rule of “Right”

- Should consider the “right”
  - Perspective
  - Outcome
  - Alternative

© Jeffrey Hoch, PhD
Mistakes

• Data
  1. Not both cost and effect
  2. Wrong cost perspective
  3. Wrong outcome
  4. Fake or wrong alternative

Main idea: Anatomy of common mistakes
Per chronic disease prevented?

Improving chronic disease prevention and screening in primary care: results of the BETTER pragmatic cluster randomized controlled trial

Eva Grunfeld1, Donna Manca2, Rahim Moineddin1, Kevin E Thorpe4,5, Jeffrey S Hoch6,7,10,11, Denise Campbell-Scherer8, Christopher Meaney4, Jess Rogers8, Jacyd Bech7, Paul Krueger11,
1) general and 2) moderate mental illness. The interventions involved a multifaceted, evidence-based, tailored practice-level intervention with a Practice Facilitator, and a patient-level intervention involving a one-hour visit with a Prevention Practitioner where patients received a tailored ‘prevention prescription’. The primary outcome was a composite Summary Quality Index of 28 evidence-based chronic disease prevention and screening actions with pre-defined targets, expressed as the ratio of eligible actions at baseline that were met at follow-up. A cost-effectiveness analysis was conducted.

Results: 789 of 1,260 (63%) eligible patients participated. On average, patients were eligible for 8.96 (SD 3.2) actions at baseline. In the adjusted analysis, control patients met 23.1% (95% CI: 19.2% to 27.1%) of target actions, compared to 28.5% (95% CI: 20.9% to 36.0%) receiving the practice-level intervention, 55.6% (95% CI: 49.0% to 62.1%) receiving the patient-level intervention, and 58.9% (95% CI: 54.7% to 63.1%) receiving both practice- and patient-level interventions (patient-level intervention versus control, P < 0.001). The benefit of the patient-level intervention was seen in both strata. The extra cost of the intervention was $26.43CAN (95% CI: $16 to $44) per additional action met.

Figure 4 Costs and effects for control and treatment groups. The ratio of the difference in costs between two groups to the difference in eligible actions accomplished represents the incremental cost-effectiveness ratio (ICER) (dotted line). PP/PF: Combined practice-level and patient-level intervention. N.E. The PF intervention is within the “efficiency frontier” so it is not considered an efficient use of resources.
Figure 4 Costs and effects for control and treatment groups. The ratio of the difference in costs between two groups to the difference in eligible actions accomplished represents the incremental cost-effectiveness ratio (ICER) (dotted line). PP/PF: Combined practice-level and patient-level intervention. NE: The PF intervention is within the "efficiency frontier" so it is not considered an efficient use of resources.
Main idea: Anatomy of common mistakes

Uncertainty

- What to do if you don't have or don't know something for your analysis?
- 2 key questions:
  1) Is it important?
     - Why not check?
  2) Would doing your “experiment” a bunch of times help?
     - To “characterize” the uncertainty
Uncertainty: Sensitivity analysis

The ROI is X or the CBA shows Y vs. based on your beliefs, this is how things could turn out...

Uncertainty

• What to do if you don't have or don't know something for your analysis?

• 2 key questions:

  1) Is it important?
     • Why not check?

  2) Would doing your “experiment” a bunch of times help?
     • To “characterize” the uncertainty
Uncertainty: Statistical analysis

95% confidence interval (or something like that)

Mistakes

- Data
  1. Not both cost and effect
  2. Wrong cost perspective
  3. Wrong outcome
  4. Fake or wrong alternative

- Analysis
  5. Not the right time horizon
  6. Not a difference ($\Delta \Delta \Delta$) or ratio of $\Delta$'s
  7. Only an estimate, no uncertainty
  8. Only 1 type of uncertainty (e.g., SA)
Main idea: Anatomy of common mistakes

Types of Mistakes

- Doing
  - Data
  - Analysis

- Using
  - Results

• Step 0: Think
• Other matters

Is what you are seeing making sense clinically?

© Jeffrey Hoch, PhD
Other matters
Not initially cost-effective ≠ don’t fund it

Can the cost-effectiveness be improved?

• Reduced price will reduce ΔC
• Targeted use will increase ΔE

Can the budget impact be improved?

Are there other factors that matter to decision makers?

© Jeffrey Hoch, PhD
Other matters

Not *initially* cost-effective ≠ don’t fund it

Can the budget impact be improved?
• Reduced price will reduce $C \times N$
• Targeted use will reduce $C \times N$

Are there other factors that matter to decision makers?
• Equity?
• Voter appeal (social pressure)?
Other matters

Not initially cost-effective ≠ don’t fund it

Can the cost-effectiveness be improved?
• Reduced price will reduce ΔC
• Targeted use will increase ΔE

Can the budget impact be improved?
• Reduced price will reduce C \times N
• Targeted use will reduce C \times N

Are there other factors that matter to decision makers?
• Equity?
• Voter appeal (social pressure)?

Special areas

• Willingness to pay more for treatments in
  ____________________________________________
  • A) Cancer
  • B) Blood Safety
  • C) Mental illness / Drug Addiction
  • D) Neonates, babies, children
  • E) Some of above

© Jeffrey Hoch, PhD
Mistakes

• Data
  1. Not both cost and effect
  2. Wrong cost perspective
  3. Wrong outcome
  4. Fake or wrong alternative

• Analysis
  5. Not the right time horizon
  6. Not a difference (Δ Δ Δ!) or ratio of Δ’s
  7. Only an estimate, no uncertainty
  8. Only 1 type of uncertainty (e.g., SA)

Using the results

9. Don’t ask, “Do the results make sense?”
10. Believe the economic results are the only thing that matters

Implications

• An economic evaluation must fill in the letters in the statement:
  • In A years, it will cost $B to get one more unit of C when using D instead of E in
    patients of type F in context G.

• Different choices for A – G create different cost-effectiveness “results”.

• When the analysis has different A – G’s from your ideal, it is problematic.

Questions?

jshoch@ucdavis.edu

Questions and answers appear on the next pages...

Q: How can I find recordings of the previous talks?

Google search

Cost NAPCRG
Questions, continued

• Q: What are some suggestions for integrating CEA into research grants?

Response

CEA can be introduced as a “third aim” into a research grant. Typically, the outcome is already decided upon (in order to calculate sample size for the grant) and the outcome data are being collected as part of the study. A key decision is whether it is worth it (and how) to collect some resource use data (e.g., hospitalizations, emergency room use and doctor visits). Once the data on cost and outcome exist, it is possible to analyze them using cost-effectiveness methods for a cost-effectiveness data set. Alternatively, data from the trial could be used to build a decision model. The decision model can extend past the trial and/or consider other outcomes or populations. For example, see http://tinyurl.com/y8hovts6 and http://tinyurl.com/y7znzhws

Also, see the two previous (referenced on the previous slide) talks for other ideas and examples.

• Q: Would you recommend any analyses that split patients/people by latent classes, so the final statement then will become several statements?

Response

Yes, I think hypothesis generation with patient subgroups is a great idea. If you are analyzing a cost-effectiveness data set, you can do stratification, add interaction terms, or use methods for latent classes. If you are making a decision model, introducing latent classes can be difficult; however, creating subgroups is not hard—make a separate model (either structure or data) for each separate group.

• As an example, Mahoney and colleagues studied the “Long-term cost-effectiveness of early and sustained clopidogrel therapy for up to 1 year in patients undergoing percutaneous coronary intervention after presenting with acute coronary syndromes without ST-segment elevation”. Her Figure 4 shows the likelihood of cost-effectiveness for 4 different groups.

Questions, continued, continued

• Q: I’m often asked to do CEA or CBA on interventions. Many times I have difficulty finding a good match for a control group. Do you have any broad guidance around matching?

Response

Yes, this can be tricky. The “real world” evidence people and the observational data crowd continue to struggle with this. I don’t feel there are any easy answers beyond trying to make the two groups comparable. One trick I sometimes use is to say, “How much would this unknown variable need to be before the new intervention is not economically attractive?” This type of threshold analysis or break even analysis can help counter the lack of a good match in control group by allowing you to see how sensitive your results are to the parameter estimate you don’t have (or do have but is not precisely estimated). Missing a good control group affects both outcomes as well as economic evaluation studies.

• Q: Could we get the slides?

Response

Certainly. I will email them to the organizers.