The Need for Improvements in Methodology to Support Evidence Development

Marc L. Berger, M.D.
Vice President
USHH Outcomes Research & Management
Merck & Co., Inc.
Methodology Meets Policy

• What makes information useful to decision makers?
  - What do we know?
    • How confident are we about what we know scientifically?
  - What would we like to know?
    • Are there critical gaps in our scientific knowledge?
      - Use of colloquial evidence as a complement or substitute for missing scientific evidence

• Options for decision makers
  - Make decisions based upon imperfect information
    • Combine scientific evidence with contextual considerations
      - Resources, Values, Habits, Traditions, Political Judgments
    • Revisit decisions if and when more information become available
  - Defer decision and ask for additional information
  - “Coverage with Evidence Development”
Evidence Review and Synthesis (Formulary Committee Staff, EPCs, DERP)

Framing Key Questions
Rigor Required

Evidence-Based Decision Making (Formulary Committee of PBM, Private Payer, States, Federal Gov’t – CMS)

Economic Information

Studies

Budget Constraints

Acceptability

Decisions

“Evidence is inherently uncertain, complex, contestable, and rarely complete.”
Transparency Requires that the Use of Different Kinds of Evidence is Clearly Identified

**Scientific Evidence**  
(ex Rigorous Evidence Reviews)  
- Context-free  
- Explicit (codified & propositional)  
- Systematic (transparent & explicit methods)  
- Replicable

**Social Science Evidence**  
(ex CEA)  
- Context-sensitive  
- Explicit  
- Systematic  
- Goes beyond Scientific Evidence

**Colloquial Evidence**  
(ex Values, Habits, Tradition, Political Judgment, Pragmatics & Contingencies, Resources, Professional Experience & Expertise, Advocacy Groups)  
“Anything that establishes a fact or gives reason for believing something”  
- Context-sensitive  
- Practical  
- Defined by relevance
Comparative Effectiveness: Asking The Right Questions, Choosing The Right Method

Methodological choice should be driven by policy goals.

by Steven M. Teutsch, Marc L. Berger, and Milton C. Weinstein

ABSTRACT: The Medicare Prescription Drug, Improvement, and Modernization Act (MMA) of 2003 has placed renewed focus on assessing the comparative effectiveness of various therapeutic options. Unfortunately, all of the evidence needed to fully assess these options is rarely available to drug formulary decisionmakers. Comparative randomized trials frequently fail to find differences when there indeed are some, while decision-modeling approaches are more likely to identify differences where there are none. We consider the consequences of these strategies. This paper proposes a framework for using different methods to assess available evidence. We contend that choosing the appropriate method can occur only when there are clear policy goals.
Policy Goals and Certainty of Comparative Evidence

What is the tolerance for error?

• **Type I** - Error of Commission (False Positive)
  – Thinking there is a difference when there is none

• **Type II** - Error of Omission (False Negative)
  – Thinking there is no difference when there is one
Level of Certainty Required Depends on the Clinical Situation

- Serious Life-threatening Health Condition
  - Greater willingness to accept less than perfect information as decisions must be made by doctors and patients

- Asymptomatic Population at Risk for Developing an Illness
  - Greater desire for certainty before widespread use
### Acting on Imperfect Evidence: How Much Regret Are We Ready to Accept?

**B. Djulbegovic et al J Clin Oncol 23: 6822-6825, 2005**

<table>
<thead>
<tr>
<th>Goals of Treatment</th>
<th>Uncertainty About Benefits/Harms of Alternative Tx Options</th>
<th>Benefit-Harm Ratio</th>
<th>Regret Associated with Wrong Recommendations</th>
<th>Acceptable Evidentiary Standards</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevention of the dz in asx individuals</td>
<td>High</td>
<td>There are important trade-offs between benefits and harms</td>
<td>High</td>
<td>Highest standard of experimental evidence</td>
<td>Tamoxifen in prevention of breast cancer</td>
</tr>
<tr>
<td>Cure</td>
<td>Low</td>
<td>The vast majority of practitioners believe that the intervention does more harm than good.</td>
<td>Low</td>
<td>May accept lower level of evidence</td>
<td>Antibiotics in H Pylori positive MALT gastric lymphoma</td>
</tr>
<tr>
<td>Prolongation of survival in years</td>
<td>Low</td>
<td>The vast majority of practitioners believe that the intervention does more harm than good.</td>
<td>Low</td>
<td>May accept lower level of evidence</td>
<td>Imatinib in chronic myeloid leukemia</td>
</tr>
<tr>
<td>Prolongation of survival in months</td>
<td>High</td>
<td>There are important trade-offs between benefits and harms</td>
<td>High</td>
<td>Highest standard of experimental evidence</td>
<td>Chemotherapy in metastatic lung cancer</td>
</tr>
<tr>
<td>Palliation (improvement in QoL)</td>
<td>Low/Moderate</td>
<td>The vast majority of practitioners believe that the intervention does more harm than good.</td>
<td>Low</td>
<td>May accept lower level of evidence</td>
<td>Morphine in pain control</td>
</tr>
<tr>
<td>Palliation (improvement in QoL)</td>
<td>High</td>
<td>It is not clear whether the intervention does more good than harm</td>
<td>Moderate</td>
<td>High quality evidence particularly if costs are high</td>
<td>Bisphosphonates in prevention of skeletal-related morbidities</td>
</tr>
</tbody>
</table>
Health Technology Assessment and Evidence-based Medicine

• Assessment of benefits, harms, and costs of alternative treatments based upon best available evidence

• Types of Evidence
  – RCTs
    • Practical Clinical Trials
  – Observational Studies
    • Longitudinal Registries
    • Case – Control, Cohort Studies, etc.
      – Retrospective Claims Analyses
      – Retrospective Analyses based on EMR or Chart Review
      – Prospective Observational Studies
  – Systematic Reviews
    • Meta-analyses
# Outstanding Methodologic Issues

<table>
<thead>
<tr>
<th>Type of Study</th>
<th>Issues</th>
</tr>
</thead>
</table>
| **RCTs**            | Generalizability  
Head-to-Head Outcomes Studies are complex, require large sample sizes over long time periods |
| **Treatment Registries** | Effectiveness:  
- Channeling and other potential biases  
- Use of surrogate vs outcomes endpoints  
Safety: Sample Size and NNH  
May lack comparison group  
Need for GCPs (AHRQ Handbook Project) |
| **Retrospective Analyses** | Effectiveness:  
- Bias (including incomplete and inaccurate information)  
- Correlation vs Causation  
Safety: Signal Identification vs Validation |
| **Systematic Reviews** | Questions addressed  
Appropriate standards of evidence  
Analysis, Interpretation and Communication of Results (Eisenberg Center) |
Good Decisions Require Good Methods