MEDICAL AND PHARMACY COVERAGE DECISION-MAKING AT THE POPULATION LEVEL

Penny Mohr
Vice President, Program Development
Center for Medical Technology Policy
OVERVIEW

• Health Spending Trends
• Current State of Evidence
• Case Study – Implantable Defibrillators
• The Promise of Comparative Effectiveness Research
• Techniques to Deal with Heterogeneity of Treatment Effects (HTE) That Would Help Decision Makers
HEALTH CARE SPENDING
US is an outlier – 30% higher than expected
 QUESTION #1: WHAT FRACTION OF INCREASING HEALTH CARE COSTS ARE RELATED TO MEDICAL TECHNOLOGY?

A. 5%
B. 25%
C. 50%
D. 80%
TECHNOLOGY AND SPENDING

- Estimates vary
  - BCBSA report: 18%
  - Project Hope: 25-33%
  - David Cutler: 50%
  - Victor Fuchs: 81% of economists identify technology as primary cost driver in health care
THE STATE OF EVIDENCE FOR DECISION-MAKING
### Evidence for purchasing goods

<table>
<thead>
<tr>
<th>Brand &amp; Model</th>
<th>Ratings and Test results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overall score</td>
</tr>
<tr>
<td>Hitachi DS18DMR</td>
<td>85</td>
</tr>
<tr>
<td>Tougher job drill/drivers</td>
<td></td>
</tr>
<tr>
<td>Makita LXT BDF451</td>
<td>82</td>
</tr>
<tr>
<td>Tougher job drill/drivers</td>
<td></td>
</tr>
<tr>
<td>Milwaukee 0824-24</td>
<td>81</td>
</tr>
<tr>
<td>Tougher job drill/drivers</td>
<td></td>
</tr>
<tr>
<td>Panasonic EY6432GQKW</td>
<td>80</td>
</tr>
<tr>
<td>General use drill/drivers</td>
<td></td>
</tr>
<tr>
<td>Bosch 33818-2G</td>
<td>80</td>
</tr>
<tr>
<td>Tougher job drill/drivers</td>
<td></td>
</tr>
<tr>
<td>Makita 6347DWDE</td>
<td>79</td>
</tr>
<tr>
<td>Tougher job drill/drivers</td>
<td></td>
</tr>
<tr>
<td>Ryobi P813</td>
<td>77</td>
</tr>
<tr>
<td>General use drill/drivers</td>
<td></td>
</tr>
<tr>
<td>Makita 6980FDWDE</td>
<td>75</td>
</tr>
<tr>
<td>Cordless impact drivers</td>
<td></td>
</tr>
<tr>
<td>Ryobi P230C</td>
<td>74</td>
</tr>
<tr>
<td>Cordless impact drivers</td>
<td></td>
</tr>
</tbody>
</table>
**EVIDENCE SUMMARY: RADIATION THERAPY FOR CLINICALLY LOCALIZED PROSTATE CANCER**

<table>
<thead>
<tr>
<th>Comparisons</th>
<th>Disease specific survival</th>
<th>Freedom from biochemical failure</th>
<th>GU/GI toxicity</th>
</tr>
</thead>
<tbody>
<tr>
<td>RT vs NT</td>
<td>insufficient</td>
<td>insufficient</td>
<td>insufficient</td>
</tr>
<tr>
<td>SBRT vs EBRT</td>
<td>insufficient</td>
<td>insufficient</td>
<td>insufficient</td>
</tr>
<tr>
<td>SBRT vs HDBRT</td>
<td>insufficient</td>
<td>insufficient</td>
<td>insufficient</td>
</tr>
<tr>
<td>SBRT vs LDBRT</td>
<td>insufficient</td>
<td>insufficient</td>
<td>insufficient</td>
</tr>
<tr>
<td>EBRT vs HDBRT</td>
<td>insufficient</td>
<td>insufficient</td>
<td>insufficient</td>
</tr>
<tr>
<td>EBRT vs LDBRT</td>
<td>insufficient</td>
<td>insufficient</td>
<td>insufficient</td>
</tr>
<tr>
<td>HDBRT vs LDBRT</td>
<td>insufficient</td>
<td>insufficient</td>
<td>insufficient</td>
</tr>
<tr>
<td>Combined mod.</td>
<td>insufficient</td>
<td>insufficient</td>
<td>insufficient</td>
</tr>
<tr>
<td>Intra SBRT</td>
<td>insufficient</td>
<td>insufficient</td>
<td>insufficient</td>
</tr>
<tr>
<td>Intra EBRT</td>
<td>insufficient</td>
<td>moderate</td>
<td>moderate</td>
</tr>
<tr>
<td>Intra LDBRT</td>
<td>insufficient</td>
<td>insufficient</td>
<td>insufficient</td>
</tr>
</tbody>
</table>

Despite lots of options, there is very little evidence!

Source: Tufts Evidence-based Practice Center: Draft AHRQ Technical Assessment, March 25, 2010
QUESTION #2: IN CARDIOLOGY, WHAT % OF CLINICAL PRACTICES ARE SUPPORTED BY HIGH QUALITY EVIDENCE

A. 10%
B. 30%
C. 50%
D. 99%
QUALITY OF EVIDENCE FOR GUIDELINE RECOMMENDATIONS IN CV DISEASE

Over 50% of evidence based on Level C evidence

THE EVIDENCE PARADOX

• 18,000+ RCTs published each year
• Tens of thousands of other clinical studies
• Systematic reviews intended to inform clinical and health policy decisions routinely conclude that evidence is inadequate
MEDICARE DECISIONS ABOUT PAYING FOR NEW TECHNOLOGY: THE CASE OF ICDs
STATUTORY BASIS FOR MEDICARE COVERAGE

• Sect. 1862 (a)(1)(A), Title 18, SSA
• “Notwithstanding any other provisions of law . . . no payment may be made... for items or services . . [which] are not reasonable and necessary for the diagnosis or treatment of illness or injury.”
• “reasonable and necessary” never defined in law or regulation
REASONABLE AND NECESSARY

• Working definition per CMS
• Adequate evidence to conclude that the item or service
  – improves net health outcomes
    • That matter to patients
  – generalizable to the Medicare population
Kaplan-Meier Survival by Treatment Group

Hazard Ratio = 0.69

Adjusted P=0.016

31% reduction in risk of all-cause mortality

Total Mortality
CONV: 19.8%
ICD: 14.2%

Hazard Ratio = 0.69
Adjusted P=0.016

Huge financial implications of the medical technology, expanding coverage to 2-3 million persons
Patients with pacemakers were excluded.
CMS analysis of the MADIT II dataset supplied by Guidant.

Subpopulation that would have greater benefit

Patients with pacemakers were excluded.
CMS analysis of the MADIT II dataset supplied by Guidant.

Subpopulation that would have greater benefit

p-value=0.25
QUESTION #3: FOR WHICH PATIENTS SHOULD MEDICARE COVER DEFIBRILLATORS?

A. Everyone with EF < 30%
B. Only people with low EF and wide QRS
C. Anybody for whom the doctor recommends one
D. Sorry, I was texting a friend, and missed the case study
No difference in some subgroups - drove home discussion how evidence should determine coverage, especially with subgroup analyses
CLINICAL EXPERTS REACT TO MEDICARE ICD POLICY

• “The Medicare program cannot prove that this technology does not provide a benefit, and therefore is obligated to pay for it.”

• “I find it hard to believe that in a country as wealthy as the US, we cannot find the funds to pay for lifesaving technology”

• “What Hitler was unable to do, the Medicare program is trying to finish”
MORTALITY BY INTENTION-TO-TREAT

Amiodarone vs. Placebo
HR: 1.06, 95% CI: 0.86, 1.30, p-value: 0.529

ICD Therapy vs. Placebo
HR: 0.77, 95% CI: 0.62, 0.96, p-value: 0.007

Sudden Cardiac Death (SCD) - HeFT (Heart Failure Trial)
**ADDITIONAL SUBGROUPS:**
**ICD vs. PLACEBO**

<table>
<thead>
<tr>
<th>Patient Group</th>
<th>N</th>
<th>HR</th>
<th>97.5% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>382</td>
<td>0.96</td>
<td>0.58, 1.61</td>
</tr>
<tr>
<td>Male</td>
<td>1294</td>
<td>0.73</td>
<td>0.57, 0.93</td>
</tr>
<tr>
<td><strong>LVEF</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 30%</td>
<td>1390</td>
<td>0.73</td>
<td>0.57, 0.92</td>
</tr>
<tr>
<td>&gt; 30%</td>
<td>285</td>
<td>1.08</td>
<td>0.57, 2.07</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 65</td>
<td>1098</td>
<td>0.68</td>
<td>0.50, 0.93</td>
</tr>
<tr>
<td>≥ 65</td>
<td>578</td>
<td>0.86</td>
<td>0.62, 1.18</td>
</tr>
<tr>
<td><strong>QRS Duration</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 120 ms</td>
<td>977</td>
<td>0.84</td>
<td>0.62, 1.14</td>
</tr>
<tr>
<td>≥ 120 ms</td>
<td>699</td>
<td>0.67</td>
<td>0.49, 0.93</td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>1283</td>
<td>0.78</td>
<td>0.61, 1.00</td>
</tr>
<tr>
<td>Non-White</td>
<td>393</td>
<td>0.75</td>
<td>0.48, 1.17</td>
</tr>
<tr>
<td><strong>Enrolling Country</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U.S.</td>
<td>1512</td>
<td>0.82</td>
<td>0.65, 1.04</td>
</tr>
<tr>
<td>Non-U.S.</td>
<td>164</td>
<td>0.37</td>
<td>0.17, 0.82</td>
</tr>
<tr>
<td><strong>Beta Blocker</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>1157</td>
<td>0.68</td>
<td>0.51, 0.91</td>
</tr>
<tr>
<td>No</td>
<td>519</td>
<td>0.92</td>
<td>0.65, 1.30</td>
</tr>
<tr>
<td><strong>Diabetes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>524</td>
<td>0.95</td>
<td>0.68, 1.33</td>
</tr>
<tr>
<td>No</td>
<td>1152</td>
<td>0.67</td>
<td>0.50, 0.90</td>
</tr>
</tbody>
</table>

- Nonsignificant difference in QRS intervals
- Findings from SCD differ from MADIT II study
- More difficult to make policy
IMPROVING EVIDENCE FOR DECISIONS
COMMON GAPS IN EVIDENCE

• Research subjects are highly selected
• Research settings are not typical of settings in which care is usually delivered
• Missing or incorrect comparators
• Physiologic or surrogate outcomes, not functional status, long term impacts, QoL

GREAT EXPECTATIONS FOR COMPARATIVE EFFECTIVENESS RESEARCH

• “Better information about the costs and benefits of different treatment options, combined with new incentive structures reflecting the information...is essential to putting the country on a sounder long-term fiscal path.”

– Peter Orszag testimony, June 2007
THE CER HYPOTHESIS

• Gaps in evidence will be reduced with greater engagement of decision makers (patients, clinicians, payers) in:
  – Deciding which questions to study
  – Working with researchers on study design
FROM COMPARATIVE EFFECTIVENESS RESEARCH (CER) TO PATIENT CENTERED OUTCOMES RESEARCH (PCOR)

• Original focus was on improving information for patients, clinicians, payers and policy makers
  – Better decisions in context of anticipated payment and delivery system reforms

• Emphasis now shifted to primary emphasis on information needs of patients, especially “patient-centered outcomes”

• But health policy forces behind original interest in CER and creation of PCOR Institute (PCORI) have not vanished
PCORI BACKGROUND AND MISSION

PCORI Mission

The Patient-Centered Outcomes Research Institute helps people make informed health care decisions – and improves health care delivery and outcomes – by producing and promoting high integrity, evidence-based information – that comes from research guided by patients, caregivers and the broader health care community

- PCORI will produce knowledge by supporting new research and the analysis and synthesis of existing research
- The statutory language defining PCORI authorizes research that supports a strong “patient-centered” orientation
Working Definition of Patient-Centered Outcomes Research

Patient-Centered Outcomes Research (PCOR) helps people make informed health care decisions and allows their voice to be heard in assessing the value of health care options. This research answers the following patient-focused questions:

1. Given my personal characteristics, conditions and preferences, what should I expect will happen to me?

2. What are my options and what are the benefits and harms of those options?

3. What can I do to improve the outcomes that are most important to me?

4. How can the health care system improve my chances of achieving the outcomes I prefer?
FROM EXPLANATORY TRIALS TO CER: COMPOUNDING HETEROGENEITY OF TREATMENT EFFECTS
Efficacy vs. Real World Effectiveness Studies
WHAT DOES HETEROGENEITY OF TREATMENT EFFECT (HTE) LOOK LIKE?

Subpopulation with skewed distribution of outcomes

Patient and treatment are the fundamental sources of heterogeneity.

All sources of HTE can interact with each other.

Segal, JB, Weiss C, Varadhan R. Understanding Heterogeneity of Treatment Effects in Pragmatic Trials With an Example of a Large, Simple Trial of a Drug Treatment for Osteoporosis. A White Paper on behalf of the Center for Medical Technology Policy
Types of HTE Analyses

Confirmatory Analysis
Test hypotheses related to subgroup effects

Exploratory Analysis
Generate hypotheses for future studies
Refine or develop predictive models of beneficial and adverse responses of individuals to treatments
METHODS TO ADDRESS HTE

- Covariate adjustment
- Pre-defined subgroup analysis
- Risk stratification
  - using validated multivariate risk prediction tool
  - run-in baseline periods or observational studies to develop propensity scores
  - consider broader psycho-social environment
SUMMARY

• Do not eliminate heterogeneity from trials

• Heterogeneity in CER can provide useful information to make informed policy and health care decisions

• Comparative effectiveness research should be designed as explorations of heterogeneity of treatment effect rather than as evaluation of average treatment effect
  – Use CER for confirmatory and exploratory analysis of subgroup effects to take advantage of the rich data obtained

• Understanding heterogeneity is a critical priority area for future research