What is Comparative Effectiveness Research?

James S. Goodwin, MD
George and Cynthia Mitchell Distinguished Chair in Geriatric Medicine
Director, Sealy Center on Aging
University of Texas Medical Branch
Types of CER

- Systematic reviews of the literature
  - Meta Analysis (two types)

- Analyses of large databases such as electronic medical records or national Medicare data

- Large randomized controlled trials that are representative of the population of real patients
Effectiveness research: strengths and weaknesses of two major approaches

<table>
<thead>
<tr>
<th>Type of Study</th>
<th>Internal validity</th>
<th>External validity</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCT</td>
<td>great</td>
<td>questionable</td>
</tr>
<tr>
<td>Observational</td>
<td>Questionable</td>
<td>Can be great</td>
</tr>
</tbody>
</table>
Assessing Toxicities of Therapies Using Observational Data

- **Advantages:**
  - Large representative sample allows estimation of toxicities in subgroups (e.g. very old, those with multiple comorbidities)
  - Ability to assess toxicity in "real world" rather than in setting of clinical trial with motivated patients, excellent monitoring, etc.
  - Ability to assess for rare toxicities and late toxicities.

- **Disadvantages:**
  - Selection bias
  - Under reporting of outcomes and covariates
  - Possible contamination by publicity (e.g. silicone breast implants)
Late toxicity: the example of anti-androgens for prostate cancer

- GnRH agonists were overused in the 1990’s and early 2000’s in men with prostate cancer

- Androgens promote bone integrity

- Does GnRH agonist use lead to increased fractures?
No Androgen Deprivation

GnRH Agonist 1-4 Doses

GnRH Agonist 5-8 Doses

GnRH Agonist ≥ 9 Doses

Orchiectomy

Table 4. Estimated Number Needed to Harm for the Occurrence of Any Fracture within 12 to 60 Months after Diagnosis According to Age and Extent of Androgen Deprivation.*

<table>
<thead>
<tr>
<th>Age</th>
<th>Gonadotropin-Releasing Hormone Agonist</th>
<th>Orchiection</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1-4 doses</td>
<td>5-8 doses</td>
</tr>
<tr>
<td>66-69 yr</td>
<td>74 (50-146)</td>
<td>42 (29-73)</td>
</tr>
<tr>
<td>70-74 yr</td>
<td>69 (46-146)</td>
<td>39 (27-71)</td>
</tr>
<tr>
<td>75-79 yr</td>
<td>61 (41-125)</td>
<td>34 (24-61)</td>
</tr>
<tr>
<td>&gt;80 yr</td>
<td>46 (32-91)</td>
<td>26 (19-45)</td>
</tr>
</tbody>
</table>

* Estimates were calculated on the basis of adjusted rates of fracture five years after diagnosis from a Cox model with any fracture as the outcome. Doses of a gonadotropin-releasing hormone agonist were grouped according to the number of doses received within the 12 months after diagnosis. CI denotes confidence interval.

Has Risk from Radiation Changed over Time as methods have improved?

Methods

• Compared women with left versus right-sided breast cancer over time

Patient Inclusion (N=27,283)

• Women in SEER (1973-2000 dataset)
• In situ, localized, or regional breast cancer
• Diagnosed between 1973 and 1989
• Known disease laterality
• Treated with primary surgical therapy followed by adjuvant radiation

(from Giordano et al., J Natl Cancer Inst 2005; 97:419-424)
Mortality from heart disease in women receiving radiation, by breast cancer laterality and year of diagnosis

(from Giordano et al., J Natl Cancer Inst, 2005; 97:419-424)
Survival free of jaw toxicity after IV bisphosphonates in 14,349 users vs 28,698 matched controls

Comparing outcomes of different treatments with observational data: fraught with difficulty

- This is a major goal of the CER initiative

- Many, many papers published comparing the results of alternative treatments based on observational data, for example, different treatments for cancer or different treatments for ischemic heart disease

- There is growing concern that such analyses cannot provide valid information because of selection bias.
Selection Bias

- Treatments in observational studies are not randomly assigned. They are chosen by patients and physicians, based on many different factors.

- Those factors may not all be measurable in the available data, yet they may influence the treatment outcome.

- The example of self-rated health.

- A variety of statistical techniques have been proposed to account for selection bias, but it is unclear if they accomplish that.
Using SEER Medicare data to examine the power of selection bias

- SEER/Medicare allows us to look at effect of various treatments for cancer on mortality

- SEER provides data on cause of death, which can serve as an internal control on selection bias

- In other words, an effective treatment for a given condition should only improve mortality from that condition, not from other conditions
The direction and effect of potential selection biases in observational studies comparing outcomes of different treatments

<table>
<thead>
<tr>
<th>Treatments compared</th>
<th>Potential for selection bias</th>
<th>Effect of bias on hazard of death</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Healthier patients</td>
<td>Worse cancer</td>
</tr>
<tr>
<td>Adjuvant Chemotherapy vs. none</td>
<td>↑↑ ↑↑</td>
<td>↑↑↑</td>
</tr>
<tr>
<td>Adjuvant radiotherapy vs. none</td>
<td>↑↑</td>
<td>↑↑↑</td>
</tr>
<tr>
<td>Surgery vs. Radiotherapy</td>
<td>↑↑↑</td>
<td>↓↓</td>
</tr>
<tr>
<td>Any treatment vs. no treatment</td>
<td>↑↑↑</td>
<td>↑↑ or ↓↓*</td>
</tr>
</tbody>
</table>

* Two types of patients might be likely to receive no treatment: those with less aggressive tumors, such as men who choose “watchful waiting” with low grade prostate cancers; and those diagnosed with highly aggressive metastatic cancers, such as multifocal glioblastoma

(from Giordano et al. Cancer, 2008; 112(11):2456-66)
An example of a bad idea:

“Survival associated with treatment vs. observation of localized prostate cancer in elderly men” (Wong et al, JAMA 2006; 296:2683-93)

- Stratification by propensity to adjust for confounders
- Treated patients had significantly better all-cause survival than untreated men (HR=0.69; 95% CI=0.60, 0.72).
Thinking about Survival after treatment for prostate cancer: a re-examination of the data

- We used identical methods, but examined effect of prostate cancer treatment on mortality for specific causes of death.

<table>
<thead>
<tr>
<th>Cause of Death</th>
<th>HR (95% CI) Associated with treatment vs. non-treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>All causes</td>
<td>0.67 (0.64; 0.70)</td>
</tr>
<tr>
<td>Prostate cancer</td>
<td>0.71 (0.59; 0.85)</td>
</tr>
<tr>
<td>All non cancer death</td>
<td>0.60 (0.50; 0.65)</td>
</tr>
<tr>
<td>Heart disease</td>
<td>0.63 (0.50; 0.69)</td>
</tr>
<tr>
<td>Diabetes</td>
<td>0.48 (0.32; 0.72)</td>
</tr>
</tbody>
</table>

(from Giordano et al., Cancer, 2008; 112 (11):2456-66)
Survival after a diagnosis of prostate cancer according to treatment and also compared to age-matched men without cancer

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Hazard of death</th>
<th>(95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non Cancer</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>With Radical Prostatectomy</td>
<td>0.60</td>
<td>(0.54; 0.66)</td>
</tr>
<tr>
<td>With Radiation</td>
<td>0.90</td>
<td>(0.85; 0.96)</td>
</tr>
<tr>
<td>Watchful waiting</td>
<td>1.27</td>
<td>(1.19; 1.36)</td>
</tr>
</tbody>
</table>

*Adjusted for age at diagnosis, SEER region, year of diagnosis, and comorbidity

(from Giordano et al. Cancer, 2008; 112(11):2454-66)
Conclusion

Comparative Effectiveness Research is a term used to describe a broad range of activities that aim to assess which individual patients might benefit from a given treatment. This is a very important goal. Some of the techniques are relatively untested and may not survive detailed scrutiny.