How Can We Use RCT Data To Assess Heterogeneity of Treatment Effects? Let Me Count The Ways

June 3, 2014
Will Patients Continue To Get the Medications They Need?

How much choice is needed?
Wide Range in Formulary Breadth

Number of Drugs Covered by Benchmark Plan

- 565-820 Drugs (54% - 84%)
- 840-960 Drugs (85% - 91.3%)
- 943-1004 Drugs (91.4% - 97.5%)
- 1007-1023 Drugs (Over 97.5%)

* Based on data released by CMS on state selected benchmark plans, February 20, 2013, available at http://cciio.cms.gov/resources/data/ehb.html. Maximum potential medicine count is 1032; totals may double-count medicines that are categorized in more than one USP class.
HTE Awareness Has Grown

AIDS advocates say drug coverage in some marketplace plans is inadequate

Ariana Eunjung Cha, December 9, 2013

“...some plans don’t cover widely used drugs. In addition, he said, some require doctors to prescribe a specified medicine before they cover others, which could compel people to take drugs more toxic to them...”

“The fear is that they are putting discriminatory plan designs into place to try to deter certain people from enrolling by not covering the medications they need...”
Not Everyone Understands HTE

MedPAC Weighs Factoring CER Into Part B Drugs, Services Payments
Lisa Gillespie, March 13, 2014, Inside CMS

“If you have two equivalent treatments, why would you want to pay more for one?”
Topics For Discussion

1. Introduction to the various approaches
   - Robert W. Dubois, MD, PhD, National Pharmaceutical Council

2. Systematic review of RCTs
   - Daniel Mullins, PhD, University of Maryland School of Pharmacy

3. Repeated period cross-over and n-of-1 studies
   - David M. Kent, MD, MSc, Tufts medical Center

4. Individual patient data and meta-analytic techniques
   - Richard Wilke, PhD, Pfizer
HTE: Biologic and Non-Biologic Factors

C. Daniel Mullins, PhD
Professor, PHSR Department
HTE: Biologic and Non-Biologic Factors

Objective: To demonstrate potential usefulness and challenges associated with identifying evidence of HTE from systematic literature reviews, using stage-4 prostate cancer as an example to explore patient-level specific HTE factors.
Conceptual Framework for Interactions and Implications of Biologic and Non-Biologic factors in Heterogeneity of Treatment Effect
Systematic Review of HTE in s4PC

• Methodology
  – MEDLINE and the PubMed electronic databases were searched for English language, human studies published between January 1946 and March 2012
  – Of the final 92 Journal articles selected
    • 87 articles studied the role of biologic factors in HTE
      – genetic factors, age, race, co-morbidities, prior treatment, clinical signs and symptoms, laboratory data and measures of s4PC disease severity
    • 5 articles studied the role of non-biologic factors in HTE
      – social, geographic and dietary factors
Systematic Review of HTE in s4PC

• Characteristics of the 92 studies
  – Conducted in the USA, Canada, or Europe
  – 16 multi-regional studies and half were multi-center studies
  – Included subgroup analyses, cohort studies, and registry data
  – Two studies were pre-specified RCTs that studied the impact of different factors on s4PC treatment-outcomes
  – Post-hoc analyses of RCTs (46%) and comparative observational studies (50%) comprise the majority of the studies
Systematic Review of HTE in s4PC

• AHRQ quality guidelines
  – Quality of the articles was rated as good, fair or poor
    • Majority were of fair quality
    • 15% were of good quality
    • 16% were of poor quality
Systematic Review Findings

• Clinical characteristics of the 92 articles
  – Most common (74%) treatment protocol was hormonal therapy
    • Androgen deprivation therapy, peripheral androgen blockade, orchiectomy and estrogen therapy
  – Main outcome in 53 (58%) was Overall Survival
    • Of these, 25 articles (27%) articles studied OS only
  – Seven percent examined HRQOL
  – Two percent examined adverse events
**Biologic**

- Biomarkers
  - HER2 expression
  - AR-CAG repeat length
  - Prostatic AR content
  - AR binding activity
  - Nuclear AR immunostaining intensity
  - Tumor growth fraction/Ki67
  - Immunostaining
  - CXCR4 expression
  - PDGFR phosphorylation
  - UPAR forms
  - TMPRSS2-ERG expression
  - Growth fraction/Ki67
  - Immunostaining
  - Tumor cellular proliferation
  - Fraction
  - Ploidy of metastases

- Race
  - Age

- Comorbidities
  - Prior treatment
  - Ischemic cerebral disease
  - Ischemic heart disease
  - Intermittent claudication
  - Decompensated heart disease
  - Venous thrombosis
  - All CV diseases
  - Concomitant diseases

- Disease severity
  - Clinical signs/symptoms
  - Grade
    - Stage
    - Gleason score
    - Visceral metastases
    - Bone scan (progression, index)
  - Extent of disease
  - Soloway score
  - Months fracture free
  - History of skeletal fracture
  - Risk group
  - Duration of disease, time to CRPC
  - Pattern of disease progression
  - Malignant pleural effusion
  - Tumor growth/regression constant
  - Liver scan
  - BM biopsy

- Laboratory data
  - PSA (Baseline, 4wk, 2m, 3m, 6m)
  - PSA velocity, rate of decrease (4wk, 12wk)
  - PSA nadir
  - Time to PSA nadir, normalization
  - PSADT (pre treatment, post nadir)
  - Time to halving time
  - PSA Response (decline, progression, % decrease)
  - Log PSA, log PSA velocity
  - Testosterone level (baseline, 3m, 6m)
  - PAP (baseline, 1m, 3m, 6m, flare)
  - ALP (baseline, 1m, 3m, 6m, flare)
  - LDH
  - CTC
  - SGOT
  - BAP
  - PRL
  - Albumin
  - ESR
  - CRP
  - BUN
  - Creatinine
  - LH
  - FSH
  - SHBG
  - CEA

- Medications
  - ADT
  - Orchiectomy
  - Flutamide
  - Estrogen therapy
  - Chemotherapy
  - Ischemic cerebral disease
  - Ischemic heart disease
  - Intermittent claudication
  - Decompensated heart disease
  - Venous thrombosis
  - All CV diseases
  - Concomitant diseases

- Non-biologic

**Diet**

- Country of residence

**Social/Behavioral**

- Social life
- Professional life
- Sexual life
- Partner status

**Geographic**

- PC-SPES

**Biologic**

- RRP
- Radiation therapy
- ADT
- Orchietomy
- Flutamide
- Estrogen therapy
- Chemotherapy

**Non-biologic**

- Pain
- Bone pain
- Performance status
  - General health status
  - Global QOL
  - Fatigue
  - Urologic symptoms
  - Days of motor deficit in MSCT
  - Ambulatory status before RT in MSCT
  - BMI

- PSA (Baseline, 4wk, 2m, 3m, 6m)
- PSA velocity, rate of decrease (4wk, 12wk)
- PSA nadir
- Time to PSA nadir, normalization
- PSADT (pre treatment, post nadir)
- Time to halving time
- PSA Response (decline, progression, % decrease)
- Log PSA, log PSA velocity
- Testosterone level (baseline, 3m, 6m)
- PAP (baseline, 1m, 3m, 6m, flare)
- ALP (baseline, 1m, 3m, 6m, flare)
- LDH
- CTC
- SGOT
- BAP
- PRL
- Albumin
- ESR
- CRP
- BUN
- Creatinine
- LH
- FSH
- SHBG
- CEA

- HgB (baseline, 3m)

* Number of good quality articles on overall survival
^ Number of good quality articles on outcomes other than overall survival
Factor explored in 3 or more articles with overall survival as the outcome
Results for HTE Factors in s4PC Patient, Outcome = OS

<table>
<thead>
<tr>
<th>HTE Factor</th>
<th>No. of times factor studied, N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Association with outcomes, direction of</td>
<td></td>
</tr>
<tr>
<td>correlation and quality of evidence*</td>
<td>Significant association</td>
</tr>
<tr>
<td></td>
<td>+ OR – correlation**</td>
</tr>
<tr>
<td></td>
<td>(Good, Fair, Poor)</td>
</tr>
<tr>
<td>Age (older age at study entry or diagnosis)</td>
<td>-2</td>
</tr>
<tr>
<td></td>
<td>(1,1,0)</td>
</tr>
<tr>
<td>Race (AA vs. non-AA)</td>
<td>-2</td>
</tr>
<tr>
<td></td>
<td>(0,2,0)</td>
</tr>
<tr>
<td>Clinical signs/symptoms</td>
<td>-28</td>
</tr>
<tr>
<td></td>
<td>(9,16,3); +2 (1,1,0)</td>
</tr>
<tr>
<td>Disease severity</td>
<td>-58</td>
</tr>
<tr>
<td></td>
<td>(11,43,4)</td>
</tr>
<tr>
<td>Gene/Biomarker expression</td>
<td>+2</td>
</tr>
<tr>
<td></td>
<td>(0,1,1); -2 (0,1,1)</td>
</tr>
<tr>
<td>Laboratory data***</td>
<td>-96</td>
</tr>
<tr>
<td></td>
<td>(20,67,9); +33 (9,22,2)</td>
</tr>
<tr>
<td>Prior treatment</td>
<td>+3</td>
</tr>
<tr>
<td></td>
<td>(3,0,0); -1 (1,0,0)</td>
</tr>
<tr>
<td>Concomitant medications</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(0,0,0)</td>
</tr>
<tr>
<td>Co-morbidities</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(0,0,0)</td>
</tr>
<tr>
<td>Social</td>
<td>+2</td>
</tr>
<tr>
<td></td>
<td>(0,2,0)</td>
</tr>
<tr>
<td>Social life****</td>
<td>+1</td>
</tr>
<tr>
<td></td>
<td>(0,1,0)</td>
</tr>
<tr>
<td>Partner status (married vs. single)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(0,0,0)</td>
</tr>
</tbody>
</table>

|                                                | No significant association   |
|                                                | (Good, Fair, Poor)           |
|                                                | Total number of times factor studied |
|                                                | (Good, Fair, Poor)           |
| Age (older age at study entry or diagnosis)    | 12                            |
|                                                | (2,9,1)                       |
| Race (AA vs. non-AA)                           | 7                             |
|                                                | (3,4,0)                       |
| Clinical signs/symptoms                        | 9                             |
|                                                | (2,5,2)                       |
| Disease severity                               | 29                            |
|                                                | (7,20,2)                      |
| Gene/Biomarker expression                      | 3                             |
|                                                | (1,2,0)                       |
| Laboratory data***                             | 52                            |
|                                                | (12,35,5)                     |
| Prior treatment                                | 5                             |
|                                                | (4,1,0)                       |
| Concomitant medications                        | 1                             |
|                                                | (0,1,0)                       |
| Co-morbidities                                 | 0                             |
|                                                | (0,0,0)                       |
| Social                                         | 0                             |
|                                                | (0,0,0)                       |
| Social life****                                | 0                             |
|                                                | (0,0,0)                       |
| Partner status (married vs. single)            | 2                             |
|                                                | (0,2,0)                       |

*Quality of evidence based on AHRQ guidelines
**+ is a positive correlation and – is a negative correlation with OS
***Elevated laboratory values
****Social life assessed by questionnaire including a score for degree of impairment of family/social life due to the medical condition or the treatment
Lessons Learned and Challenges for HTE in Other Diseases

• Searching for HTE requires casting a wide net
  – HTE not a reliable search term
  – HTE goes by many names (some incorrect)

• Pre-specified HTE factors aid literature searches
  – Adds to the number of articles
  – BUT you don’t know what you don’t know

• HTE challenging to detect retrospectively
  – Analytic rigor of empirical analysis
  – Interpretation of study reports
Lessons Learned and Challenges for HTE in Other Diseases

- HTE likely to be a secondary not primary aim
  - Best case is pre-specified secondary aim
  - Post hoc analyses have methodological challenges

- Many studies provide subgroup analyses
  - Not all address HTE
  - Not always clear how to disentangle principal effect from interaction of \{HTE factor x treatment\}
Lessons Learned and Challenges for HTE in Other Diseases

• As a proportion of the vast literature in s4PC, focus on HTE is small
  – Most s4PC HTE literature on biologic v. non-biologic
  – Existing evidence leaves significant gaps regarding HTE

• Literature that addresses HTE may be less extensive for other disease states than s4PC
  – Need more studies
  – Need appropriate study designs to address HTE
Variation in individual treatment effects from repeated period cross-over and multi-person N-of-1 studies
Distinguish between “person-level” heterogeneity of treatment effect (HTE) and “group-level” HTE.

Review why commonly used approaches for group-level HTE are poor surrogates for person-level HTE.

Present state of the literature of repeated period cross-over studies.
Clinical Trial Results

**Individual Treatment Effects in a Deterministic Framework: Four possibilities**

<table>
<thead>
<tr>
<th>Subject Name</th>
<th>Without Treatment</th>
<th>With Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAM</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>MARY</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>BOB</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>BEN</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CHRISTINE</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NEIL</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>MOHAMED</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>JENNIFER</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PAUL</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NISHA</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>MIGUEL</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>LAYLA</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PAUL</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>EMANUEL</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CHERYL</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PATRICK</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>OSCAR</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>JULIANNE</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>THOMAS</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GEORGE</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

\[0 = \text{alive} \quad 1 = \text{dead}\]
# Clinical Trial Results with Counterfactuals

<table>
<thead>
<tr>
<th>Subject Name</th>
<th>Without Treatment</th>
<th>With Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAM</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>MARY</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>BOB</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>BEN</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>CHRISTINE</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>NEIL</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>MOHAMED</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>JENNIFER</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PAUL</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>NISHA</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>MIGUEL</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>LAYLA</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>PAUL</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>EMANUEL</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>CHERYL</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PATRICK</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>OSCAR</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>JULIANNE</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>THOMAS</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GEORGE</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

**HARM**

**BENEFIT**

### Individual Treatment Effects in a Deterministic Framework: Four possibilities

<table>
<thead>
<tr>
<th>Without Treatment</th>
<th>With Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

0 = alive

1 = dead
# Clinical Trial

<table>
<thead>
<tr>
<th>Without Treatment</th>
<th>With Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Proportion

<table>
<thead>
<tr>
<th>Dead</th>
<th>Without Treatment</th>
<th>With Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>11/20</td>
<td>55%</td>
<td>9/20</td>
</tr>
</tbody>
</table>

**BENEFIT**
Why do we fail to reliably detect HTE?

1. Information failure
   - Observable co-variates are totally unrelated to the causal determinants of HTE.
   - The causal mechanisms may be so complex that HTE cannot be distinguished from a stochastic/probabilistic process.

2. Analytic failure
   - Low power
   - Limitations of conventional (one-variable-at-a-time) subgroup analysis
Aim

- Provide a broad summary of all repeated period cross over studies
Methods

- Conduct an electronic search on Medline®, EMBASE™, Web of Science, and Cochrane Central for all published randomized trials that utilize repeated cross-over randomized trials including multi-person n-of-1 trials designs examining pharmacological agents.

- Repeated period cross-over randomized trials: those in which patients are randomized to sequences of treatments with at least one treatment received in more than one period.
Evidence Map: Methods

- Describe the studies.
- Summarize descriptions and claims of HTE.
N-of-1 Studies: Flow Diagram

N-of-1 citations identified by recent search (2011 to Feb 2014) (n=2676)

Articles retrieved for full-text review (n=48)

Abstracts did not meet criteria (n=2613)

Articles did not meet criteria (n=97)
Reasons:
- Articles with single patient (n= 41)
- Non-pharmacological trial (n= 12)
- Reviews (n=29)
- Not relevant study design (n=6)
- Abstracts and protocol (n=9)

Articles identified by a prior systematic review [Gabler et al 2011] (n =100)

Total eligible N-of-1 articles (N=51)
Repeated Period Crossover Studies: Flow Diagram

Repeat cross-over citations identified by searches in MEDLINE and bibliography search (Through March 2014) (n=10596)

 Articles retrieved for full-text review (n=325)

Abstracts that did not meet eligibility criteria (n=10271)
Reject reasons: non-pharmacological trials; reviews; irrelevant study design

Full-texts that did not have repeated cross-over design (n=257)

Full texts accepted for repeated cross-over design (N=68)
# N-of-1 Study Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>N studies</td>
<td>51</td>
</tr>
<tr>
<td>Publication Years</td>
<td>1986 - 2014</td>
</tr>
<tr>
<td>Population</td>
<td>Pediatric 12 studies, Adults 39 studies, Elderly 15 studies</td>
</tr>
<tr>
<td>Disease Conditions (Top 5)</td>
<td>ADHD 7 studies, Obstructive Airway disease 6 studies, Osteoarthritis 6 studies, Chronic pain 5 studies, GERD 5 studies</td>
</tr>
<tr>
<td>Major Systems studied</td>
<td>Neurology 13 studies, Arthritis/Rheumatology 10 studies, Psychiatry 9 studies, Gastrointestinal 7 studies, Respiratory 6 studies, Miscellaneous 6 studies, Sleep disorders, Allergy, Cancer, Muscular, Vascular</td>
</tr>
<tr>
<td>N subjects enrolled</td>
<td>Total 1974, Mean 39 (SD 61), Median 21 (IQR 10 – 43)</td>
</tr>
<tr>
<td>N subjects completed</td>
<td>Total 1573, Mean 32 (SD 52), Median 14 (IQR 14 – 34)</td>
</tr>
<tr>
<td>Comparisons</td>
<td>Active drug 9 studies, Placebo 41 studies, Active drug and placebo 1 study</td>
</tr>
<tr>
<td>Most frequently studied primary outcome</td>
<td>Self-reported outcomes 15 studies, Symptom improvement 13 studies, Behavior 8 studies</td>
</tr>
</tbody>
</table>
# Repeated Period Crossover Study Characteristics

<table>
<thead>
<tr>
<th>N studies</th>
<th>68</th>
</tr>
</thead>
<tbody>
<tr>
<td>Publication Years</td>
<td>1976 - 2012</td>
</tr>
</tbody>
</table>

### Disease Conditions (Top 5)
- Conditions
  - Angina: 11 studies
  - HTN: 8 studies
  - HA/Migraine: 5 studies
  - GERD: 4 studies
  - Pain: 4 studies
  - Miscellaneous (36 studies)
  CTD, Depression, Dysmenorrhea, ADD/ADHD, ED, Enuresis, Others (1 each; 20 studies)

### Major Systems studied
- Organ systems
  - Cardiovascular: 11 studies
  - Neurology: 11 studies
  - Psychiatry: 9 studies
  - HTN: 8 studies
  - GI: 5 studies
  - Miscellaneous (24 studies)
  - Pain, Urology, GYN, Rheumatology, Heme/Onc, Allergy, Dermatology, Drug abuse, Endocrine, Lipids, Nephrology, Ophthalmology, Respiratory

### N subjects enrolled in 51 studies
- Total 3753
- Mean 58 (SD 94)
- Median 32 (IQR 21 – 52)

### N subjects completed
- Total 2295
- Mean 20 (SD 57)
- Median 24 (IQR 13 – 40)

### Comparison (n studies)
- Active drug only: 16 studies
- Placebo only: 45 studies
- Active drug and placebo: 7 studies
Description of HTE: Evidence map
Conclusions
Thank you!
Individual patient data and meta-analytic techniques: Some alternative methods

Richard J. Willke, PhD, Pfizer

For presentation at Workshop 11:

HOW CAN WE USE RANDOMIZED TRIAL DATA TO ASSESS HETEROGENEITY OF TREATMENT EFFECTS? LET ME COUNT THE WAYS

ISPOR 19th Annual International Meetings, Montreal, June 3, 2014

GLOBAL HEALTH & VALUE
Introduction

• Heterogeneity of treatment effect (HTE) is an increasingly important consideration in the development and evaluation of medicines, including HTA and economic evaluation

• Traditional subgroup analysis relies primarily on stratification by individual covariates, but there are a variety of other methods available

• I will briefly cover:
  • An overview of such methods
  • A meta-analytic method: Model-based meta-analysis
  • Two outcomes-based stratification methods:
    • Quantile Regression
    • Finite Mixture Modeling
Mixed coverage decisions by NICE

An Analysis of NICE’s ‘Restricted’ (or Optimized’) Decisions
– Phill O’Neill and Nancy J. Devlin. Pharmacoeconomics 2010

• Between 2006-2009, over half of NICE decisions (69) for medicines were “mixed” – neither “yes” or “no” for all patients
• Of the 69 mixed decisions, the majority recommended coverage for less than half of patients who would have been eligible based on the license
• Often these decisions are made based on cost-effectiveness results for patient subgroups
• “NICE HTA no. 171, Lenalidomide for the Treatment of Multiple Myeloma in People Who Have Received at Least One Prior Therapy.[5] The guidance recommended lenalidomide as an option for people who have received two or more prior therapies. The licence for this medicine, as stated in the ‘Technology’ section of NICE’s guidance, states that it is licensed for use in people who have received one prior therapy.”
A Partial List of Methods*

• Subgroup analysis of clinical trials
• Meta-analysis & meta-regression
• Predictive risk modeling
• Classification and Regression Tree Analysis
• Latent growth, finite mixture and growth mixture modeling
• Series of n-of-1 trials
• Quantile regression
• Non-parametric methods

### Characteristics of various methods

<table>
<thead>
<tr>
<th>Intent of HTE Analysis</th>
<th>Meta-analysis</th>
<th>CART</th>
<th>N of 1 trials</th>
<th>LGM/FMM/GMM *</th>
<th>QTE**</th>
<th>Non-parametric</th>
<th>Predictive risk models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exploratory and confirmatory</td>
<td>Exploratory</td>
<td>Exploratory and initial testing</td>
<td>Exploratory, initial testing, and confirmatory</td>
<td>Exploratory, initial testing, and confirmatory</td>
<td>Exploratory and confirmatory</td>
<td>Initial testing and confirmatory</td>
<td></td>
</tr>
<tr>
<td>Data Structure</td>
<td>Trial summary results, possibly with subgroup results</td>
<td>Panel or cross-section</td>
<td>Repeated measures for a single patient: time series</td>
<td>Time series and panel</td>
<td>Panel and cross-sectional</td>
<td>Panel, time series, and cross-sectional</td>
<td>Panel or cross-sectional</td>
</tr>
<tr>
<td>Data Size Considerations</td>
<td>Advantage of combining small sample sizes</td>
<td>Large sample sizes</td>
<td>Small sample sizes</td>
<td>Small sample: LGM/FMM</td>
<td>Moderate to large sample sizes</td>
<td>Large sample sizes</td>
<td>Sample sizes depends on specific risk function</td>
</tr>
<tr>
<td>Key Strength(s)</td>
<td>Increase statistical power by pooling of results</td>
<td>Doesn’t require assumptions around normality</td>
<td>Patient is own control</td>
<td>Accounting for unobserved characteristics</td>
<td>Robust to outcome outliers</td>
<td>No functional form assumptions</td>
<td>Multivariate approach to identifying risk factors or HTE</td>
</tr>
<tr>
<td>Key Limitation(s)</td>
<td>Included studies must be similar</td>
<td>Fairly sensitive to changes in underlying data</td>
<td>Requires de novo study</td>
<td>Criteria for optimization solutions not clear</td>
<td>Treatment effect designed for a quantile, not a specific patient</td>
<td>Computationally demanding</td>
<td>May be more or less interpretable or useful clinically</td>
</tr>
</tbody>
</table>

* LGM/FMM/GMM: Latent growth modeling/Finite mixture modeling/Growth mixture modeling; **QTE: Quantile treatment effect
Determine the level of existing evidence of HTE

Step 1

- Strong
- Mixed Results
- Weak or No prior beliefs

Step 2

- Consider HTE study design
- Treatment(s)
- Intent of HTE Analysis
- Source(s) of HTE
- Study Features

Step 3

Choose the most appropriate analysis method

- CART
- LGM/GMM
- Meta
- N of 1
- NP
- Predictive
- QTE
- ...

• Exploratory Studies for HTE
• Initial Testing for HTE
• Confirmatory Research for HTE

• Data Structure (cross-sectional, time series, and panel data)
• Data Size Considerations
• Key Strength(s) and Limitation(s)
Model-based meta-analysis

• A variety of meta-regression that assumes the trial outcome is a specific function of trial-level covariates (i.e., “the model”)*
  – Meta-regression typically assumes linear models, while model-based meta-analysis often uses non-linear models
  – Otherwise, literature review and analysis standards are the same as other types of meta-analysis

• Increasingly used by Pharmacometrics (clinical pharmacology modeling) to estimate relationships between safety and efficacy outcomes and different doses of a treatment
Figure 1  Relationship between predicted (solid line) and observed (symbols) placebo-adjusted percentage change from baseline in low-density lipoprotein (LDL) cholesterol after fibrate treatment and the baseline level of triglycerides. Each symbol in the graph represents one trial, with the size of the symbol proportional to the precision of the observation.

* From Mandema et al, Clinical Pharmacology & Therapeutics (2011)
Uses of MBMA in Outcomes Research

1. As an estimation option for any health economic variable that one might analyze using meta-regression, due to the desire to control for selected sample characteristics

2. As an alternative approach for conducting indirect or mixed treatment comparisons
   - E.g., as a robustness check on network meta-analysis

3. As an input to health economic models to
   - Translate surrogate marker effects to clinical outcomes
   - Allow for differing assumptions around patient populations, if used as inclusion/exclusion criteria for trials, to modify treatment effects*

*see Slejko et al, 2014.
Outcome-based stratification

• Simple comparison of outcomes distributions by treatment group
  – Ignores predictive factor differences; possibly subject to other biases

• Quantile regression (QR)
  – Allows for conditioning of outcomes on observed predictive factors
  – Treatment and covariate effects estimated at each quantile
  – Absolute deviation approach less sensitive to outliers than OLS

• Finite mixture modeling (FMM)
  – Allows data to determine natural outcomes “clusters”
  – Clustering and effects within clusters can be conditioned on predictive factors
  – Estimated marginal effects take cluster probabilities into account
• Provides additional distributional information about the central tendency and statistical dispersion of the treatment effect in a population.

Suppose the $\tau^{th}$ conditional quantile function is:

$$Q_{Y|X}(\tau) = X\beta_\tau$$

Given the distribution function of $Y$, $\beta_\tau$ can be obtained by solving:

$$\beta_\tau = \arg\min_{\beta \in \mathbb{R}^k} E(\rho_\tau(Y - X\beta)).$$

Solving the sample analog gives the estimator of $\beta$:

$$\hat{\beta}_\tau = \arg\min_{\beta \in \mathbb{R}^k} \sum_{i=1}^{n}(\rho_\tau(Y_i - X\beta)).$$
Quantile regression example

- Taken from (Bitler et al, What Mean Impacts Miss: Distributional Effects of Welfare Reform Experiments. AmEconRev, Sept 2006)
- Previous evaluations of welfare/jobs programs primarily focused on mean impacts; theory gives ambiguous predictions about expected impact
- Mean impacts can miss important distributional differences in impacts of such programs
- The Connecticut Jobs First program allowed for a 21-month “disregard” of all earned income (up to poverty line) while on welfare, much more than on standard welfare, encouraging recipients to get “startup” jobs
- Quantile regression allowed for a more detailed analysis of how the Jobs First program affect recipients’ earnings
Quantile Regression results for Jobs First program

- Difference in earnings by earnings quantile, Jobs First program*
- Shows no impact at low or high end, significant impact in 55-80^{th} percentiles

* Bitler et al, Am Econ Rev, 2006
Finite Mixture Model (FMM) basics

The density function for a $C$-component finite mixture is

$$f(y|x; \theta_1, \theta_2, ..., \theta_C; \pi_1, \pi_2, ..., \pi_C) = \sum_{j=1}^{C} \pi_j f_j(y|x; \theta_j)$$

where $0 < \pi_j < 1$, and $\sum_{j=1}^{C} \pi_j = 1$

Estimation can be carried out using maximum likelihood

$$\max_{\pi, \theta} \ln L = \sum_{i=1}^{N} \left( \log \left( \sum_{j=1}^{C} \pi_j f_j(y|\theta_j) \right) \right)$$
Finite mixture model example

• Taken from Deb et al, “The effect of job loss on overweight and drinking,” _J Health Econ_ 2013.
• Previous studies had given varying and ambiguous results, partially attributed to heterogeneity of effects
• Used data on job loss associated with business closures (presumed to be exogenous job loss) and personal characteristics, from Health & Retirement Survey
• Found two component mixture when BMI was the outcome variable; being younger, depressed, lower income predicted presence in the “second” component.
FMM results for BMI*

* Reproduced from Deb et al, 2013

Robust standard errors in parentheses. Regressions also include year dummies.

\[ \pi_1, p < 0.05 \]

\[ \pi_1, p < 0.01 \]
Final thoughts

• Most of these methods are available in the major statistical packages (SAS, Stata, R)

• They won’t all give the same results; a little work to understand how the varying methods and their assumptions affect the results can be very informative

• Given the trends in both payer decision-making and personalized medicine, prospective evaluation of heterogeneity of treatment effect for new medical treatments is quickly becoming a necessity
References