Designing an Open-Science Pediatric Learning Health System

Presented at the NIH/AHRQ Conference on Methodological Challenges in Comparative Effectiveness Research, Bethesda, MD December 3, 2010

Christopher B. Forrest, MD, PhD
Professor of Pediatrics
Children’s Hospital of Philadelphia
University of Pennsylvania School of Medicine
Email: forrestc@email.chop.edu
Acknowledgment

The technical components of this work are being funded by a grant from the Agency for Healthcare Research and Quality:

AHRQ R01HS020024

The learning community building (patient/family engagement and patient/provider collaboration) components are funded by a TR01 grant from the National Institutes of Health/NIDDK:

R01 DK085719

These two grants represent a $20 million investment by the federal government.
Goal & Objectives of this Project

Goal
Develop a prototype for a learning health system in pediatrics focusing on patients with inflammatory bowel disease.

Objectives
1. EHR-based registry;
   “Data-in-once ➔ Information-out-many-times”
2. Quality: population and self management
3. CER: timing of introduction of biologics
4. Sustainability, spread, and governance
Open-Science Principles for the Learning Health System

Open source software, open notebook, open access to data, open access publications

Benefits

Rapid feedback, collaboration, repurposing, accelerating impact
Improving Outcomes with a Learning Health System

- Patients and Families
- Clinicians
- Identify Uncertain Management Practices
- Comparative Effectiveness Research
- Electronic Health Records
- Registry Database
- Registry Applications
- Distributed Network of Registries
- Standardize Process
  - Reduce Variability in Process
  - Customize Process to Patient Needs
- Patient Outcomes

New Knowledge

- Identify New Gaps in Care

Enhanced Registry
The Learning Community

- Pediatric EHR Data Sharing Network (PEDSNet)
  - 15 Children’s Hospitals; 8 participating
- ImproveCareNow – IBD collaborative
ImproveCareNow: 29 Pediatric GI Practice Sites, from small private groups to large Academic Medical Centers, 270 physicians, and 10,000 children with IBD
Percent of IBD Patients in Remission (PGA)
Focus on the Outcome
Question
What is the main patient outcome that we are trying to achieve?

Answer
remission = inactive disease
Patients in remission >70%

- Accurate diagnosis and disease classification
- Appropriate drug selection
- Appropriate drug dosing
- Timing of Biologic Agents – an IOM Top 100 CER Priority
- Adequate nutritional intake
- Appropriate growth monitoring

- Registry
- Population Management Tool
- Pre-visit Planning
- Protocols
- Auditing
- Self-Management Support
Core Data-Set

Process
Learning community identifies the minimum data-set required to support improvement and analytics

Early stages
Led by clinicians

More mature stages
Full community participates in core data selection
Pediatric IBD Registry Current State: Data Recorded Three Times

- Encounter
- Abstraction
- Data Entry
Data-In-Once Requires Changes to the EHR User Interface
Pediatric IBD Registry Future State: Data-in-Once, Automated ETL Processes, Federated Queries

Sites that do not have an EHR or cannot install the i2b2 middleware will be able to (1) upload reports from local EHR or registry or (2) perform manual entry into central node.
Five Adaptive Treatment Strategies for Children with Newly Diagnosed Crohn’s Disease

- **Top-Down Biologics**: Anti-TNFα at diagnosis
- **Rapid Step-Up to Biologics**: Thiopurines/Steroids → Anti-TNFα within 6 months
- **Late Step-Up to Biologics**: Thiopurines/Steroids → Anti-TNFα 6-12 months
- **Thiopurines Only**: Thiopurines/Steroids only
- **No Biologics/Thiopurines**: Steroids only or no systemic therapy
Modified Delphi: Clinical Factors that Affect Choice to Use Biologics

<table>
<thead>
<tr>
<th>Clinical Factor</th>
<th>Mean Importance Rating (1-10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fistulizing/penetrating disease</td>
<td>9.5</td>
</tr>
<tr>
<td>Disease severity/activity</td>
<td>9.0</td>
</tr>
<tr>
<td>Failure to respond to TPs/Steroids</td>
<td>9.0</td>
</tr>
<tr>
<td>Perianal disease, including fistulas</td>
<td>8.8</td>
</tr>
<tr>
<td>Steroid dependence/resistance</td>
<td>7.5</td>
</tr>
<tr>
<td>SEs of other medications</td>
<td>6.9</td>
</tr>
<tr>
<td>Growth failure</td>
<td>6.9</td>
</tr>
<tr>
<td>Extensive small bowel disease</td>
<td>6.4</td>
</tr>
<tr>
<td>Stricturing disease</td>
<td>6.2</td>
</tr>
</tbody>
</table>
Adding Patient Reported Outcomes

• Adding a PRO measurement System using NIH’s PROMIS Network
• Patients/Parents will go to web site to complete PROMIS computerized adaptive tests before/after and encounter
• The LHS community will design this system
Accelerating Knowledge Impact with a Learning Health System

- Identify Uncertain Management Practices
- Improvement
- Change Data Collection System
- Interpretation of Results
- Analysis
Dynamic Treatment Strategies in Crohn’s Disease: Real-Time Evaluation

Presented at the NIH/AHRQ Conference on Methodological Challenges in Comparative Effectiveness Research, Bethesda, MD December 3, 2010

Marshall M. Joffe
Associate Professor of Biostatistics
Center for Clinical Epidemiology and Biostatistics
Department of Biostatistics and Epidemiology
University of Pennsylvania School of Medicine
mjoffe@mail.med.upenn.edu
Outline

• Problems:
  – Defining questions of clinical interest
  – Confounding by variables affected by treatment
  – Performing analyses in real time
Questions of clinical interest

• Strategies for using biologics for children with Crohn’s disease
  – Key question: how best to time initiation of biologics
    • Strategies defined in terms of timing
    • Compare strategies
    • Find optimal strategy

• Ideally, randomized trial comparing all strategies of interest
  – Expensive
  – Time-consuming
Types of strategies

- Top-down biologics:
  - Anti-TNFα at diagnosis
- Rapid step-up to biologics
  - Anti-TNFα within 6 months
- Late step-up to biologics
  - Anti-TNFα 6-12 months
- Thiopurines only
- No biologics/thiopurines
- Decision on treatment course specified
  - Strategies do not adapt to clinical course
Adaptive/dynamic strategies

• Allow treatment decisions to depend on clinical course
  – e.g., start anti-TNFα if current therapies appear to have failed
    • Use disease activity scale as indicator/trigger
    • Options
      – If disease activity above threshold level, initiate anti-TNF
      – If disease activity rises given amount above baseline, initiate

• Analogous to protocols in randomized trials
Comparison of approaches

• Standard
  – Compare people following both strategies
  – Possibly control for confounders at some point
  – Subjects completing strategy unrepresentative of population

• New
  – Compare what would happen to same people under both approaches

People following

<table>
<thead>
<tr>
<th>Strategy 1</th>
<th>Strategy 2</th>
</tr>
</thead>
</table>

If everyone follows

<table>
<thead>
<tr>
<th>Strategy 1</th>
<th>Strategy 2</th>
</tr>
</thead>
</table>
Evaluation/estimation

• Use marginal structural models/inverse probability weighting
• Controls for confounding of effect of treatment strategy by variables affected by strategy
  - e.g., disease activity at intermediate time points
• Conventional methods (e.g., regression) biased whether or not control for intermediate variable
• Explain graphically
Graph of setting

- Interested in effect of strategy
- Characteristics of covariate ($L_1$)
  - 1 or 1’: independent predictor of outcome
  - 2: influences subsequent treatment
  - 3: influenced by prior treatment
- 2 and (1 or 1’): covariate confounder
  - Must control
- 3 and 1: covariate intermediate
  - Should not control

$A_0, A_1$: treatment at times 0, 1
$L_1$: covariate at time 1
$U$: other covariate
$Y$: outcome
Conditions for Valid Estimation

• No unmeasured confounding variables
  – Requires measurement of variables which
    • Influence treatment decisions AND
    • Independent predictors of outcome
  – Integrated system assists with this
    • Ask physicians about what influences treatment decisions
      – Already done: Modified Delphi
      – Can be updated
    • Make certain that information on determinants completely collected at every encounter
      – Don’t unnecessarily dichotomize continuous covariates during data collection
        » May interfere with ability to control confounding
Conditions for Valid Estimation (2)

• Positive probability of staying on regime given covariate history
  – If strategy never used for subset of patient experience, cannot evaluate
    • Examples
      – If anti-TNFα never given at diagnosis for patients with low disease activity, cannot evaluate regime: “Top-down biologics”
      – If anti-TNFα never given when PCDAI<10, cannot evaluate regime: “Initiate biologics as soon as PCDAI exceeds 5”
  – Don’t want to include variables that are extremely good predictors of treatment but do not predict outcome
    • e.g., suppose that physician determines treatment given covariate history, not associated with outcome
    • Would not want to adjust for physician as confounder
Inverse probability weighting

- Idea: each subject who stays on regime at time $t$ represents
  - Self; and
  - Subjects otherwise comparable who did not stay on regime
- Mimics sequential randomized trial
  - Covariates do not affect treatment
  - Effect of treatment unconfounded
  - Apply standard methods in weighted pseudopopulation
Steps for implementing analysis

- Fit model for (repeated) treatment at $t$ as function of covariate history, prior treatment history
- Compute probability of remaining on regime at $t$ given history
- Multiply probabilities together for all times
- Weights: inverse of product of probabilities
- Create weighted pseudopopulation
- Use weighted version of standard approaches for estimating expected outcome under regime
- Compare outcomes under different regimes
Dynamic regimes: evaluation

• Several differences:
  – Individual subject’s observed treatment consistent with multiple regimes
    • e.g., subject who has PCDAI<10, no biologics throughout follow-up; consistent with
      – “No biologics”
      – “Initiate therapy when PCDAI > 10”
  – “Clone” subject: contributes as many observations as the number of regimes consistent with treatment history
    • Above example: subject contributes to both regimes
  – Account necessary for variance of contrasts
Dynamic regimes: evaluation (2)

- Link between covariate history and subsequent treatment
  - Not broken
  - Changed
    - Deterministic function of history
Evaluation in real time

• Challenge:
  – Analyses using inverse probability weighting, marginal structural models more difficult, time-consuming than standard analyses

• Speed up analyses after first one
  – Perhaps use same predictors, treatment in subsequent evaluation
  – Consider automated algorithms/machine learning, especially for treatment assignment models

• Feedback within system
  – Regimes feasible at outset may become infeasible to evaluate later as practice patterns change
  – Consider evaluating new aspects/dimensions of treatment choice over time
Feedback/Quality Improvement

• Standardization of treatment decisions: goal/danger
  – Goal: after analysis, want to avoid treatment decisions determined to be suboptimal
    • More uncertainty here than in randomized trials, because of uncertainty of ability to completely control confounding
  – Danger: do not want to completely standardize decisionmaking
    • Makes it difficult to compare standardized protocol with other possibilities within database
  – Middle course
    • Identify suboptimal regimes as result of analysis; discourage use (perhaps with automated implementation)
    • Do not attempt to standardize within region of clinical equipoise
References
