Pragmatic Trials and the Learning Health Care System: Lumbar Imaging with Reporting of Epidemiology (LIRE) Trial

Seattle Symposium
Sept 29, 2014

Patrick J. Heagerty
Department of Biostatistics, University of Washington
Fred Hutchinson Cancer Research Center
Group Health Research Institute
Acknowledgements

• AHRQ: R01HS019222-01; 1R01HS022972-01
  • BOLD / BOLDER
• NIH: UH2 AT007766-01
  • LIRE and NIH Collaboratory
• NIH: UL1 TR000423
  • WWAMI CTSA Award
Pragmatic Trials

• dealing with things sensibly and realistically in a way that is based on practical rather than theoretical considerations

• From Greek *pragmatikos* 'relating to fact’

Oxford Dictionaries

http://www.oxforddictionaries.com/us/definition/american_english/pragmatic
When To Get Pragmatic

Josephine P. Briggs, M.D.
Director
National Center for Complementary and Alternative Medicine

View Dr. Briggs' biographical sketch

The focus of this activity is **pragmatic** trials. What makes a pragmatic trial, well ... pragmatic? Pragmatic trials, sometimes called effectiveness trials, test an intervention, or compare several interventions, delivered under conditions as close to the “real world” as possible. In contrast, efficacy studies typically test interventions under ideal conditions, with very close monitoring.

The interventions tested in a pragmatic trial should not require a complex structure for implementation. The study design should incorporate rigorous controls, prospectively identified and preferably randomized. The monitoring and outcome assessment should be as simple as possible and potentially use electronic health records. And finally, the research question to be answered should be important—to patients, payers, and health care delivery systems.
Challenge: Clinical research is not relevant to practice

- Traditional RCTs study effectiveness of txs for carefully selected populations under ideal conditions.
- Difficult to translate to real world.
- When implemented into everyday clinical practice, often see a “voltage drop” — dramatic decrease in effectiveness.

“If we want more evidence-based practice, we need more practice-based evidence.”

Learning health care systems

**EVALUATE**
Collect data and analyze results to show what works and what doesn’t.

**ADJUST**
Use evidence to influence continual improvement.

**IMPLEMENT**
Apply plan in pilot and control settings.

**DESIGN**
Design care and evaluation based on evidence generated here and elsewhere.

**DISSEMINATE**
Share results to improve care for everyone.

**INTERNAL AND EXTERNAL SCAN**
Identify problems and potentially innovative solutions.

In a learning health care system, research influences practice and practice influences research.
Key features of most PCTs

**Use of electronic health records (EHRs)**

- EHRs allow efficient and cost-effective, recruitment, participant communication & monitoring, data collection, and follow up

**Randomization at clinic or provider level**

- Protocols can be tailored to local sites and can adapt to changes in a dynamic health care environment

---

NIH Collaboratory
Health Care Systems Research Collaboratory
Rethinking Clinical Trials

Supported by the Common Fund at the National Institutes of Health, the Health Care Systems Research Collaboratory is intended to improve the way clinical trials are conducted by creating a new infrastructure for collaborative research. The ultimate goal is to ensure that healthcare providers and patients can make decisions based on the best available clinical evidence.

The NIH HCS Research Collaboratory includes a Coordinating Center that provides national leadership and technical expertise in all aspects of research with healthcare systems. The Coordinating Center will make data, tools, and resources from these projects available to the greater research community to facilitate a broadened base of research partnerships with health care systems.

The NIH HCS Research Collaboratory also supports the design and rapid execution of several high-impact Pragmatic Clinical Trial Demonstration Projects (listed below) that will address questions of major public health importance that engage health care delivery systems in research partnership.

<table>
<thead>
<tr>
<th>Principal Investigator</th>
<th>Institution</th>
<th>Project</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The Collaboratory

- Supported by The Common Fund (NIH Director’s fund)
- Goal: improve the way (pragmatic) clinical trials conducted
- Build infrastructure for collaborative research
<table>
<thead>
<tr>
<th>Principal Investigator</th>
<th>Institution</th>
<th>Project</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gloria Coronado</td>
<td>Kaiser Foundation Research Institute</td>
<td>Strategies and Opportunities to Stop Colon Cancer in Priority Populations</td>
</tr>
<tr>
<td>Lynn DeBar</td>
<td>Kaiser Foundation Research Institute</td>
<td>Collaborative Care for Chronic Pain in Primary Care</td>
</tr>
<tr>
<td>Laura Dember</td>
<td>University of Pennsylvania</td>
<td>Pragmatic Trials in Maintenance Hemodialysis</td>
</tr>
<tr>
<td>Susan Huang</td>
<td>University of California--Irvine</td>
<td>Decreasing Bioburden to Reduce Healthcare–Associated Infections and Readmissions</td>
</tr>
<tr>
<td>Jeffrey Jarvik</td>
<td>University of Washington</td>
<td>A Pragmatic Trial of Lumbar Image Reporting with Epidemiology (LIRE)</td>
</tr>
<tr>
<td>Gary Rosenthal</td>
<td>University of Iowa</td>
<td>Nighttime Dosing of Anti–Hypertensive Medications: A Pragmatic Clinical Clinical Trial</td>
</tr>
<tr>
<td>Gregory Simon</td>
<td>Group Health Cooperative</td>
<td>Pragmatic trial of population–based programs to prevent suicide attempt</td>
</tr>
</tbody>
</table>
LIRE (To Read)- A Pragmatic Cluster Randomized Trial
LIRE: Background and Rationale

• Lumbar spine imaging frequently reveals incidental findings
• These findings may have an adverse effect on:
  – Subsequent healthcare utilization
  – Patient health related quality of life
<table>
<thead>
<tr>
<th>Modality</th>
<th>Author/Year</th>
<th>Age Range</th>
<th>Prev</th>
</tr>
</thead>
<tbody>
<tr>
<td>MR</td>
<td>Boden/1990</td>
<td>20-60</td>
<td>44% 93%</td>
</tr>
<tr>
<td>MR</td>
<td>Stadnik/1998</td>
<td>17-60</td>
<td>52% 80%</td>
</tr>
<tr>
<td>MR</td>
<td>Weishaupt/1998</td>
<td>20-50</td>
<td>72-100%</td>
</tr>
<tr>
<td>MR</td>
<td>Jarvik/2001</td>
<td>35-70</td>
<td>91%</td>
</tr>
</tbody>
</table>
Disc Degeneration in Asx
Disc degeneration: Approximately 80%-100% of people without back pain have this, so finding may not be related to patient’s pain.
Lumbar Spine Macro

The following findings are so common in people without low back pain that while we report their presence, they may have nothing to do with a patient’s low back pain (Reference-Jarvik et al, Spine 2001):

Finding (prevalence in pts without low back pain)
- Disc degeneration (91%)
- Disc signal Loss (83%)
- Disc height loss (56%)
- Disc bulge (64%)
- Disc protrusion (32%)
- Annular fissure (38%)
Hypothesis

• The benchmark information will influence subsequent management of primary care patients with LBP
  – Fewer subsequent imaging tests
  – Fewer referrals for minimally invasive pain treatment
  – Fewer referrals to surgery
  – Less narcotic use
LIRE UH3 Hypothesis

• For patients referred from primary care, inserting epidemiological evidence in lumbar spine imaging reports will reduce subsequent diagnostic and therapeutic interventions, including cross-sectional imaging (MR/CT), opioid prescriptions, spinal injections and surgery.
Primary Care Clinics With LBP Patients

Randomize Clinics

- Macro with Epi Info
  - Outcomes Assessment

- No Macro with Epi Info
  - Outcomes Assessment
Stepped Wedge Design

- Exposed to LIRE intervention
- Unexposed to LIRE intervention

*Randomization
wave 1
wave 2
wave 3
wave 4
wave 5

Follow-up period
Accrual period

Q1 Q2 Q3 Q4 Q1 Q2 Q3 Q4 Q1 Q2 Q3 Q4 Q1 Q2 Q3 Q4
Year 1 Year 2 Year 3 Year 4 Year 5
Stepped Wedge Design

• Staggered cluster crossover design
• Temporally spaces the intervention
• Assures that each participating clinic eventually receives the intervention
• Analysis: multilevel, longitudinal
Stepped Wedge Power

What is new?

- Approaches to power calculations for cluster randomized stepped wedge designs have been provided, but a simple sample size formula is lacking. Therefore, we present a sample size formula for these kinds of trials.
- We derived a formula in which, besides the cluster size and intracluster correlation, the number of steps and measurements can be varied.
- The stepped wedge design requires a substantial smaller sample size than a parallel group or analysis of covariance design.

Figure 5. Screenshot of the steppedwedge dialog box: Clusters tab—set up for example 1
Advantages of SW Design

• Controls for external temporal trends
• Assures all sites receive intervention
• Participation more palatable for interventions viewed as desirable
• Issue: heterogeneity of effects
The Intervention: GHC Test Template

<table>
<thead>
<tr>
<th>X-RAY SPINE LUMBAR 2V OR 3V</th>
<th>Status: Final result  MyChart: Not Released</th>
</tr>
</thead>
<tbody>
<tr>
<td>(STD) AP+LAT+LAT-L5</td>
<td>Next appt with me: None  Dx: Sprain of lumbar region</td>
</tr>
</tbody>
</table>

### Details

<table>
<thead>
<tr>
<th>Narrative</th>
</tr>
</thead>
<tbody>
<tr>
<td>[HST]: Pain upper lumbar spine</td>
</tr>
<tr>
<td>[SAS]: fell</td>
</tr>
<tr>
<td>RIS 10.7</td>
</tr>
<tr>
<td>Integration Testing Script #1</td>
</tr>
<tr>
<td>Exam and Procedure: 9631 04/11/2013 10:55:00LUMB2 (Site1) : LUMBAR 3 VIEW AP, LAT, L5</td>
</tr>
<tr>
<td>MRN: 03000045</td>
</tr>
<tr>
<td>PATIENT LAST NAME: BOOR</td>
</tr>
</tbody>
</table>

**History:** Pain upper lumbar spine
**Additional Comments:**
**Accession Notes:** fell
**Pain upper lumbar spine**

---

**Signed by:** RAD3 TST, MD  
Date: 04/11/2013  
Time: 12:10

---

**Specimen Collected:** 04/11/13 10:55 AM  
**Last Resulted:** 04/11/13 12:09 PM

### Comment

The following findings are so common in normal, pain-free volunteers that while we report their presence, they must be interpreted with caution and in the context of the clinical situation. Among people between the age of 40 and 60 years who do not have back pain, a plain film x-ray will find that about:

- 8 in 10 have disk degeneration
- 6 in 10 have disk height loss

Note that even 3 in 10 means that the finding is quite common in people without back pain.
Intervention Text

The following findings are so common in normal, pain-free volunteers that while we report their presence, they must be interpreted with caution and in the context of the clinical situation. Among people between the age of 40 and 60 years who do not have back pain, a plain film x-ray will find that about:

- 8 in 10 have disk degeneration
- 6 in 10 have disk height loss

Note that even 3 in 10 means that the finding is quite common in people without back pain.
LIRE- The Outcome

- A composite measure of spine intervention intensity—a single metric of overall intensity of resource utilization for spine care
- Passively collected from EHR
- Convert CPT codes to RVUs as our primary metric of back-related utilization
<table>
<thead>
<tr>
<th>HCPCS</th>
<th>DESCRIPTION</th>
<th>TOTAL 2012 RVUs</th>
</tr>
</thead>
<tbody>
<tr>
<td>72100</td>
<td>X-ray exam of lower spine</td>
<td>1.07</td>
</tr>
<tr>
<td>99214</td>
<td>Office/outpatient visit level 4</td>
<td>2.26</td>
</tr>
<tr>
<td>72131</td>
<td>CT lumbar spine w/o dye</td>
<td>6.27</td>
</tr>
<tr>
<td>72148</td>
<td>MRI lumbar spine w/o dye</td>
<td>11.31</td>
</tr>
<tr>
<td>63047</td>
<td>Laminectomy</td>
<td>32.89</td>
</tr>
</tbody>
</table>
### Table 5: Participating Health Care Systems

<table>
<thead>
<tr>
<th>Name</th>
<th>Location</th>
<th># of Primary Care Clinics</th>
<th># of PCPs*</th>
<th># of Enrollees</th>
<th># of Enrollees Making Low Back Pain Visits in 2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kaiser Permanente Northern California</td>
<td>Northern California</td>
<td>17</td>
<td>1,096</td>
<td>2,430,972</td>
<td>149,329</td>
</tr>
<tr>
<td>Henry Ford Health System</td>
<td>Michigan</td>
<td>26</td>
<td>230</td>
<td>187,201</td>
<td>23,877</td>
</tr>
<tr>
<td>Group Health Cooperative of Puget Sound</td>
<td>Washington State</td>
<td>24</td>
<td>303</td>
<td>347,250</td>
<td>37,675</td>
</tr>
<tr>
<td>Mayo Health System</td>
<td>Minnesota, Wisconsin</td>
<td>61</td>
<td>269</td>
<td>1,500,000 (est.)</td>
<td>106,700 (est.)</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td></td>
<td><strong>128</strong></td>
<td><strong>1,898</strong></td>
<td><strong>4,465,423</strong></td>
<td><strong>317,581</strong></td>
</tr>
</tbody>
</table>

*Primary Care Physicians*
Challenges

• System differences will always be present in large pragmatic trials

• Effectiveness of intervention within diagnosis subgroups
  – central stenosis, herniated disc, foraminal stenosis, etc.
  – Abstract radiology reports to classify patients
    • NLP
    • Amazon mTurk
    • Manual abstraction
System Heterogeneity

One-stage penalized model_CCS-S

Groups
- sig. (139)
- 0.01 - sig (144)
- 0.05 - 0.01 (147)
- > 0.05 (1755)

Show names (2185)

Search...
Heterogeneity: Example

- Backpain Outcomes using Longitudinal Data (BOLD)
- Prospective cohort of >65 with new visit for back pain
- Henry Ford / Kaiser NoCal
- 967 / 3164 subjects
Heterogeneity: Example

• Q: does the endorsement of specific CPT codes vary by site?
• Many with sparse counts
• CPT code groupings via CCS
  – Clinical Classifications Software
Heterogeneity: Example

<table>
<thead>
<tr>
<th></th>
<th>CPT$_1$</th>
<th>CPT$_2$</th>
<th>$\ldots$</th>
<th>CPT$_s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_1$</td>
<td>0</td>
<td>2</td>
<td>$\ldots$</td>
<td>0</td>
</tr>
<tr>
<td>$p_2$</td>
<td>5</td>
<td>10</td>
<td>$\ldots$</td>
<td>3</td>
</tr>
<tr>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
</tr>
<tr>
<td>$p_n$</td>
<td>1</td>
<td>0</td>
<td>$\ldots$</td>
<td>1</td>
</tr>
<tr>
<td>$Y_{ij1}$</td>
<td>$Y_{ij2}$</td>
<td>$\ldots$</td>
<td>$Y_{ij_s}$</td>
<td></td>
</tr>
</tbody>
</table>
Heterogeneity: Example

\[ \log(E(Y_i)) = \beta_0 + \beta_1 \text{Site} + \sum_{j=2}^{200} \beta_j \text{Block}_j + \sum_{j=2}^{200} \gamma_j \text{Block}_j \ast \text{Site} + \sum_{j=2}^{200} \sum_{k=2}^{n_k} \eta_{j,k} \text{CPT}_{j,k} + \sum_{j=2}^{200} \sum_{k=2}^{n_k} \theta_{j,k} \text{CPT}_{j,k} \ast \text{Site} + \log(N) \]

penalty = \lambda(\sum_j |\beta|^2 + \sum_j |\gamma|^2 + \sum_j \sum_k |\eta|^2 + \sum_j \sum_k |\theta|^2)
Heterogeneity: Example

\[ \lambda_{i,0}, \lambda_{i,1}, \lambda_{i,2}, \ldots, \lambda_{i,K} \]

\[ X_i \rightarrow Y_{i,11}, Y_{i,12}, \ldots, Y_{i,Ks} \]
Heterogeneity: Example

- Site logRR = $\beta_1 + \gamma_j + \theta_{jk}$
- L2 Penalized composite likelihood
- Dynamic graphics
- CPT code specific inference (not model-based)
System Heterogeneity
Summary: Pragmatic Trials

• Challenges in the design of pragmatic trials
• Heterogeneity / Data quality
• Integration of research and care (LIRE: data push, pull)
• QI / Care delivery
Key People

**UW**
- Jerry Jarvik, MD - Radiology
- Katie James, PA-C, MPH - Project Director
- Bryan Comstock, MS - Biostats
- Nina Pashova, PhD - Biostat
- Xu Shi, PhD student - Biostat
- Kari Stephens, PhD - Medical Informatics

**Non-UW**
- Rick Deyo, MD, MPH - OHSU
- Dan Cherkin, PhD - GHRI
- Rene Hawkes - GHRI
- Safwan Halabi, MD - HFHS
- Dave Neren, PhD - HFHS
- Dave Kallmes, MD - Mayo
- Jyoti Pathak, PhD - Mayo
- Patrick Luetmer, MD - Mayo
- Andy Avins, MD, MPH - KPNC
Thank you
CPT Code Inference
Converting CPTs to RVUs

- Validate CPT conversion by directly pulling RVUs from one site.
Challenges

• CPT counts differ by site
  – Step wedge design helps to address this since before-after comparison is within site
  – Using only back-related RVUs improves accuracy/reliability → using algorithm developed by Martin et al at Dartmouth

• Different pharmacy data systems (e.g. not all sites have Rx filled data)
  – Within-system comparisons will be valid
A pragmatic–explanatory continuum indicator summary (PRECIS): a tool to help trial designers

Kevin E. Thorpe MMath, Merrick Zwarenstein MD MSc, Andrew D. Oxman MD, Shaun Treweek BSc PhD, Curt D. Furberg MD PhD, Douglas G. Altman DSc, Sean Tunis MD MSc, Eduardo Bergel PhD, Ian Harvey MB PhD, David J. Magid MD MPH, Kalipso Chalkidou MD PhD

Published at www.cmaj.ca on Apr. 16, 2009. An abridged version of this article appeared in the May 12 issue of CMAJ. This article was published simultaneously in the May 2009 issue of the Journal of Clinical Epidemiology (www.jclinepi.com).
“Classic” Clinical Trial Business Model

Size
- Mostly small N
- Huge budgets

Endpoints
- Mostly surrogate
- Clinical trials employ adjudication

Setting
- Research enterprise — “parallel universe”
- “High-grade” data — audited, monitored

Califf RM et al. JAMA 2012;307:1838-47
Challenge: Clinical research is slow

• Traditional RCTs are slow and expensive—and rarely produce findings that are easily put into practice.

• In fact, it takes an average of 17 years before research findings lead to widespread changes in care.
Challenge: The evidence paradox

- >18,000 RCTs published each year—plus tens of thousands of other clinical studies.
- Yet systematic reviews consistently find not enough evidence to effectively inform clinical decisions providers and patients must make.
Pragmatic Trials - Model

- Size: Large n → precise estimates, evaluate heterogeneity
- Endpoints: patient oriented usually with minimal adjudication
- Setting: integrated into real world
  - Non-academic centers
  - Leverage digital data
  - Patients as partners
NIH Health Care System Collaboratory

- Collaboratory Coordinating Center
- Nighttime Dose of Anti-Hypertensive Medications
- Prevent Suicide Attempt
- Reduce Mortality in End Stage Renal Disease (sites to be selected from units across all 50 states)
- Stop Colon Cancer in Priority Populations
- Chronic Pain in Primary Care
- Reduce Infections and Readmissions
- Lumbar Image Reporting and Epidemiology

The Collaboratory

Additional sites to be determined
PCOR

- Outcomes Research
- Health Services Research
- Effectiveness Research
- Cost-effectiveness Research
- Patient Centered Outcomes Research
Patient-Centered Outcomes Research Institute: PCORI

- Independent organization
- Goal to help patients, clinicians, purchasers and policy makers make better informed health decisions
- Spearheading CER
PCOR Trust Fund

- 2010-2012: $210 million
- 2013: ~$320 million
  - $150 million general revenues
  - $1/Medicare beneficiary + private plans
- 2014-2019: ~$650 million/yr
  - $150 million general revenues
  - $2/Medicare beneficiary + private plans
PCORI Mission

PCORI 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 National Priorities

• Comparative Assessments of Prevention, Diagnosis, and Treatment Options
• Improving Healthcare Systems
• Addressing Disparities
• Accelerating Patient-Centered and Methodological Research
First Targeted Topics

• Treatment Options for Uterine Fibroids
• Treatment Options for Severe Asthma in African-Americans and Hispanics/Latinos
• Preventing Injuries from Falls in the Elderly
• Treatment Options for Back Pain
• Obesity Treatment Options in Diverse Populations

While it says tx, dx included as well
<table>
<thead>
<tr>
<th>Stakeholder Engagement Essential</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Patients</td>
</tr>
<tr>
<td>• Patient-advocacy groups</td>
</tr>
<tr>
<td>• Community groups</td>
</tr>
<tr>
<td>• Caregivers</td>
</tr>
<tr>
<td>• Health care</td>
</tr>
<tr>
<td>providers</td>
</tr>
<tr>
<td>• Professional associations</td>
</tr>
<tr>
<td>• Payers</td>
</tr>
<tr>
<td>• Industry</td>
</tr>
</tbody>
</table>
What Is Stakeholder Engagement?

- Participation in formulation of research questions
- Defining essential characteristics of study participants, comparators, and outcomes
- Monitoring of study conduct/progress
- Dissemination of research results
More Than Just Researchers

Stakeholders

Patients & Caregivers

Researchers

PCORI

Patient-Centered Outcomes Research Institute
PCORI Review Criteria

1. Impact of condition on health of individuals and populations
2. Innovation and potential for improvement through research
3. Patient-Centeredness
4. Rigorous research methods
5. Inclusiveness of different populations
6. Team and Environment
7. Efficient use of resources
<table>
<thead>
<tr>
<th>Year</th>
<th>Policy / Plan Year Ending</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>On or after October 1, 2012 and before October 1, 2013</td>
<td>$1.00 x average number of covered lives</td>
</tr>
<tr>
<td>2</td>
<td>On or after October 1, 2013 and before October 1, 2014</td>
<td>$2.00 x average number of covered lives</td>
</tr>
<tr>
<td>3 and subsequent</td>
<td>On or after October 1, 2014 and before October 1, 2019</td>
<td>($2.00 x index*) x average number of covered lives</td>
</tr>
<tr>
<td></td>
<td></td>
<td>*Index is based on increases in the projected <em>per capita</em> amount of National Health Expenditures</td>
</tr>
</tbody>
</table>
Pilot Results: Subsequent Imaging Within 1 Yr (retrospective pilot)

1/71

\[ p = 0.14 \]

\[ \text{OR}^* = 0.22 \]

* Adjusted for imaging severity
Results: Subsequent Narcotic Rx Within 1 Yr (retrospective pilot)

\[ \text{p} = 0.01 \]
\[ \text{OR}^* = 0.29 \]

\[ \frac{5}{71} = 7.0\% \]
\[ \frac{37}{166} = 22.2\% \]
Figure 1: The blank “wheel” of the pragmatic–explanatory continuum indicator summary (PRECIS) tool. “E” represents the “explanatory” end of the pragmatic–explanatory continuum.