Pragmatic Trials
Large and Small
Steps Towards a Learning Healthcare System

Data Science Course 2016
Mark J. Pletcher, MD MPH
Department of Epidemiology and Biostatistics
3 Related Concepts

• Pragmatic vs. Explanatory Trials
• Large Simple Trials
• Learning Healthcare System
Outline

• The Rationale
  • Make trials more efficient
  • Make trials more informative for decision-making

• The Concepts
  • Pragmatic vs. Explanatory, Large Simple Trials
  • 4 Examples: GISSI, TASTE, ADAPTABLE, VEST

• The Vision: Learning Healthcare System
  • Local efforts (UCSF CTSI)
  • National efforts and resources

• Science of using electronic health record data
NIH Appropriation in Current and Constant Dollars

- 1995: $11,300
- 2003: $21,003
- 2009: $35,745
- 2010: $36,209
- 2013: $30,702

- With Supplemental Appropriation (ARRA)
- Current $ (Millions)
- 1995 Constant $ (Millions)
### TABLE 3-2  Breakdown of the Costs for a Large, Global Clinical Trial (14,000 patients, 300 sites)

<table>
<thead>
<tr>
<th>Expense</th>
<th>Cost (in millions of $)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site payments</td>
<td>150.0</td>
</tr>
<tr>
<td>Monitoring</td>
<td>90.0</td>
</tr>
<tr>
<td>Data management and statistics</td>
<td>12.0</td>
</tr>
<tr>
<td>Project and clinical leadership</td>
<td>12.0</td>
</tr>
<tr>
<td>Interactive voice response systems (IVRS) and drug distribution</td>
<td>10.8</td>
</tr>
<tr>
<td>Publications</td>
<td>.1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>~300</td>
</tr>
</tbody>
</table>

**SOURCE:** Califf, 2009

From: 3. Challenges in Clinical Research


**Copyright © 2010, National Academy of Sciences.**

NCBI Bookshelf. A service of the National Library of Medicine, National Institutes of Health.
**TABLE 3-2** Breakdown of the Costs for a Large, Global Clinical Trial (14,000 patients, 300 sites)

<table>
<thead>
<tr>
<th>Expense</th>
<th>Cost (in millions of $)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site payments</td>
<td>150.0</td>
</tr>
<tr>
<td>Monitoring</td>
<td>90.0</td>
</tr>
<tr>
<td>Data management and statistics</td>
<td>12.0</td>
</tr>
<tr>
<td>Project and clinical leadership</td>
<td>12.0</td>
</tr>
<tr>
<td>Interactive voice response systems (IVRS) and drug distribution</td>
<td>10.8</td>
</tr>
<tr>
<td>Publications</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>~300</strong></td>
</tr>
</tbody>
</table>

SOURCE: *Califf, 2009.*

From: *3. Challenges in Clinical Research*


This is crazy!
Do randomized trials help us make decisions?

Traditional RCT design:

• Control conditions carefully
• Optimize internal validity and causal inference
• Help us learn biology
Do randomized trials help us make decisions?

Traditional RCT design:

• Control conditions carefully
• Optimize internal validity and causal inference
• Help us learn biology

These design decisions may not be optimal for informing decision-making
Do randomized trials help us make decisions?

Explanatory vs. Pragmatic Trials

- **Explanatory**: designed to optimize causal inference and help us learn biology

- **Pragmatic**: designed to help us compare effectiveness and make real-world decisions

Do randomized trials help us make decisions?

Explanatory vs. Pragmatic Trials

- **Explanatory:** designed to optimize causal inference and help us learn biology → laboratory conditions

- **Pragmatic:** designed to help us compare effectiveness and make real-world decisions → normal conditions

Explanatory vs. Pragmatic

Inclusion and Exclusion criteria

• Explanatory: ?

• Pragmatic: ?
Explanatory vs. Pragmatic

Inclusion and Exclusion criteria

• **Explanatory**: Carefully restricted to individuals thought likely to adhere to and respond to intervention

• **Pragmatic**: All-comers
Explanatory vs. Pragmatic

Intervention protocol

• Explanatory: ?

• Pragmatic: ?
Explanatory vs. Pragmatic

Intervention protocol

• **Explanatory**: Strict, carefully-monitored fidelity, only use experienced clinicians, avoid co-interventions

• **Pragmatic**: Flexible and practical
Explanatory vs. Pragmatic

Control

• Explanatory: ?

• Pragmatic: ?
Explanatory vs. Pragmatic

Control

• **Explanatory**: Placebo or other designed to test a specific physiologic hypothesis

• **Pragmatic**: Best alternative treatment or “usual care”
Explanatory vs. Pragmatic

Adherence

• Explanatory: ?

• Pragmatic: ?
Explanatory vs. Pragmatic Adherence

- **Explanatory**: Run-in period to weed out non-adherers, careful monitoring, “as treated” analysis

- **Pragmatic**: All-comers, see what happens, imperfect adherence is part of the intervention
Explanatory vs. Pragmatic

Measurements

• Explanatory: ?

• Pragmatic: ?
Explanatory vs. Pragmatic

Measurements

- **Explanatory**: Measure as much as possible, try to define mechanisms

- **Pragmatic**: Leave participants alone as much as possible*

* Subgroup analyses still useful
Explanatory vs. Pragmatic

Primary Outcome Measurement

• Explanatory: Measure what’s important for assessing physiology

• Pragmatic: Measure what’s clinically apparent (easier and more important to patients?)
Explanatory vs. Pragmatic

The Pragmatic-Explanatory Continuum Indicator Summary (PRECIS)

Measures pragmatism in 10 key domains

Figure 2

**a** PRECIS summary of a randomized controlled trial of self-supervised and directly observed treatment of tuberculosis (DOT) [9]

**b** PRECIS summary of the North American Symptomatic Carotid Endarterectomy Trial (NASCET) of carotid endarterectomy for symptomatic patients with high-grade carotid stenosis [10]

**c** PRECIS summary of a randomized trial of low-dose aspirin for the prevention and treatment of pre-eclampsia (CLASP) [11]

**d** PRECIS summary of a randomized trial of low-dose aspirin for the prevention of pre-eclampsia in women at high risk [12]
C

PRECIS summary of a randomized trial of low-dose aspirin for the prevention and treatment of pre-eclampsia (CLASP) [11]
Explanatory vs. Pragmatic

• Which kind of trial is cheaper?
Explanatory vs. Pragmatic

• Which kind of trial is cheaper?
  • Many “pragmatic” design features are also easier to implement

• “Large Simple Trials” – related concept that emphasizes the cost benefits
  • See JAMA Viewpoint by Eapen/Lauer/Temple
  • Strongly supported by FDA (Rob Califf→Commissioner!)

Large simple pragmatic trials

• Win-win: both cheaper and better for informing real-world clinical decisions?
Example 1: GISSI

Gruppo Italiano per lo Studio della Streptochinasi nell’infarto Miocardico (GISSI) - 1986

• RCT comparing streptokinase to usual care for all-cause in-hospital mortality in acute MI
  • 11,806 randomized, 11,712 follow-up
  • 10.7% vs 13%, RR=0.81, p<.0002
  • More benefit if given earlier

Example 1: GISSI

Gruppo Italiano per lo Studio della Streptochinasi nell’infarto Miocardico (GISSI) - 1986

• RCT comparing streptokinase to usual care for all-cause in-hospital mortality in acute MI
  • 11,806 randomized, 11,712 follow-up
  • 10.7% vs 13%, RR=0.81, p<.0002
  • More benefit if given earlier

Super-simple, large, very important results

Example 2: TASTE Trial

Thrombus Aspiration in ST-Elevation Myocardial Infarction in Scandinavia (TASTE trial)

• Registry in Sweden for angiography and angioplasty already collecting outcomes
• RCT within registry patients: thrombus aspiration vs. conventional PCI
• All-cause mortality at 30 days: No difference

Example 2: TASTE Trial

**Thrombus Aspiration in ST-Elevation Myocardial Infarction in Scandinavia (TASTE trial)**

- Registry in Sweden for angiography and angioplasty already collecting outcomes
- RCT within registry patients: thrombus aspiration vs. conventional PCI
- All-cause mortality at 30 days: No difference

**Embedded “Registry Trial”**

Example 2: TASTE Trial

**Thrombus Aspiration in ST-Elevation Myocardial Infarction in Scandinavia (TASTE trial)**

- Registry in Sweden for angiography and angioplasty already collecting outcomes
- RCT within registry patients: thrombus aspiration vs. conventional PCI
- All-cause mortality at 30 days: No difference

Cost?

Example 2: TASTE Trial

Thrombus Aspiration in ST-Elevation Myocardial Infarction in Scandinavia (TASTE trial)

• Registry in Sweden for angiography and angioplasty already collecting outcomes
• RCT within registry patients: thrombus aspiration vs. conventional PCI
• All-cause mortality at 30 days: No difference

Cost = $300,000 (compare with $300M)

Example 3: ADAPTABLE Trial

Aspirin Dosing: A Patient-Centric Trial Assessing Benefits and Long-Term Effectiveness (ADAPTABLE)

• RCT aspirin 325 vs. 81 mg for secondary prevention
• Demonstration trial for PCORnet
• Use electronic health records to:
  • Find patients (n=10,000)
  • Measure outcomes (MI, bleeding, death)

Example 3: ADAPTABLE Trial

Aspirin Dosing: A Patient-Centric Trial Assessing Benefits and Long-Term Effectiveness (ADAPTABLE)

- RCT aspirin 325 vs. 81 mg for secondary prevention
- Demonstration trial for PCORnet
- Use electronic health records to:
  - Find patients (n=10,000)
  - Measure outcomes (MI, bleeding, death)

Embedded in EHR, network

Example 3: ADAPTABLE Trial

Aspirin Dosing: A Patient-Centric Trial Assessing Benefits and Long-Term Effectiveness (ADAPTABLE)

• RCT aspirin 325 vs. 81 mg for secondary prevention
• Demonstration trial for PCORnet
• Use electronic health records to:
  • Find patients (n=10,000)
  • Measure outcomes (MI, bleeding, death)

Cost = ~$10 million

http://pcornet.org/
Example 3: ADAPTABLE Trial

Aspirin Dosing: A Patient-Centric Trial Assessing Benefits and Long-Term Effectiveness (ADAPTABLE)

- RCT aspirin 325 vs. 81 mg for secondary prevention
- Demonstration trial for PCORnet
- Use electronic health records to:
  - Find patients (n=10,000)
  - Measure outcomes (MI, bleeding, death)

Will it succeed? Is it the right model?

Example 4: VEST

The Vest Prevention of Early Sudden Death Trial

- RCT LifeVest (wearable defibrillator vest) vs. usual care → sudden death within the first 3 months for immediately post-MI patients with EF<35%.

https://clinicaltrials.gov/ct2/show/NCT01446965
Example 4: VEST

The Vest Prevention of Early Sudden Death Trial

- RCT LifeVest (wearable defibrillator vest) vs. usual care → sudden death within the first 3 months for immediately post-MI patients with EF<35%.

How pragmatic is it? Redesign ideas?

https://clinicaltrials.gov/ct2/show/NCT01446965
## Example 4: VEST

### The Vest Prevention of Early Sudden Death Trial

<table>
<thead>
<tr>
<th>Subjects</th>
<th>EF&lt;35%, no renal failure, nursing home, etc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recruitment</td>
<td>Find post MI patients in hospital, recruit and randomize</td>
</tr>
<tr>
<td>Intervention</td>
<td>Wear the Lifevest as much as possible; avoid co-interventions</td>
</tr>
<tr>
<td>Control</td>
<td>Usual care, co-interventions carefully monitored</td>
</tr>
<tr>
<td>Adherence</td>
<td>Carefully monitored, per-protocol analysis planned</td>
</tr>
<tr>
<td>Outcome</td>
<td>Sudden death, not all-cause mortality</td>
</tr>
</tbody>
</table>

[https://clinicaltrials.gov/ct2/show/NCT01446965](https://clinicaltrials.gov/ct2/show/NCT01446965)
### V2 (Month 3): Procedure Checklist

**Study ID#** | **Acrostic** | **Date Visit Commenced** | **Staff ID #**
--- | --- | --- | ---
| | | | |

#### Case Report Form# / Procedure

<table>
<thead>
<tr>
<th>Case Report Form# / Procedure</th>
<th>Please mark if done</th>
</tr>
</thead>
<tbody>
<tr>
<td>#93: Vital Status</td>
<td>Yes</td>
</tr>
<tr>
<td>[NOTE: If the participant is found to be deceased, complete CRFs #93, 45, 35, and 36. If vital status cannot be ascertained (unknown), the participant is considered &quot;lost to follow-up&quot;. Complete CRFs #93 and 45. For deceased or lost to follow-up participants, the other CRFs on this checklist do not need to be completed - Mark &quot;No, other reason (or N/A)&quot;. ]</td>
<td></td>
</tr>
<tr>
<td>#16: Medical Care Utilization</td>
<td></td>
</tr>
<tr>
<td>#17b: VEST F/U Symptoms Checklist</td>
<td></td>
</tr>
<tr>
<td>#19: F/U Medication Inventory</td>
<td></td>
</tr>
<tr>
<td>#22: LifeVest Use &amp; Satisfaction</td>
<td></td>
</tr>
<tr>
<td>#99: Device Follow-Up</td>
<td></td>
</tr>
<tr>
<td>#33: Health Questionnaire</td>
<td></td>
</tr>
<tr>
<td>#8: Echocardiogram</td>
<td></td>
</tr>
<tr>
<td>#45: Close-Out Form (if applicable)</td>
<td></td>
</tr>
</tbody>
</table>
Example 4: VEST

How could VEST have been more pragmatic?

https://clinicaltrials.gov/ct2/show/NCT01446965
Example 4: VEST

How could VEST have been more pragmatic?

• VEST prescription vs. usual care → all-cause mortality using death index, no other measurements
  • No adherence monitoring except what is clinically appropriate and feasible
  • Allow co-interventions
  • No adjudication of sudden death; all-cause mortality is what counts

https://clinicaltrials.gov/ct2/show/NCT01446965
Example 4: VEST

How could VEST have been more pragmatic?

- VEST prescription vs. usual care → all-cause mortality using death index, no other measurements
  - No adherence monitoring except what is clinically appropriate and feasible
  - Allow co-interventions
  - No adjudication of sudden death; all-cause mortality is what counts

Hard to give nothing...

https://clinicaltrials.gov/ct2/show/NCT01446965
Example 4: VEST

How could VEST have been more pragmatic?

• Could we randomize patients BEFORE we approach them?
  • Find patients via EHR search and alert system
  • Randomize to:
    • A: Approach, offer and encourage LifeVest use
    • B: Don’t approach at all – PURE Usual care (might use LifeVest)
  • Note: randomization occurs BEFORE consent

https://clinicaltrials.gov/ct2/show/NCT01446965
Example 4: VEST

How could VEST have been more pragmatic?

• Could we randomize patients BEFORE we approach them?
  • Find patients via EHR search and alert system
  • Randomize to:
    • A: Approach, offer and encourage LifeVest use
    • B: Don’t approach at all – PURE Usual care (might use LifeVest)
  • Note: randomization occurs BEFORE consent

Is this ethical?

https://clinicaltrials.gov/ct2/show/NCT01446965
Example 4: VEST

“Randomized Encouragement Design”

• Randomizing to encouragement to do something
  • Likely much lower adherence rates \(ightarrow\) much larger sample size needed
    • Adherence 40% vs. 10% instead of 90% vs. 3%?
  • But much easier??
• Works best if:
  • You are collecting outcome data \textit{passively} (e.g., via EHR)
  • You don’t need consent before randomizing
  • The intervention could be considered “quality improvement”? 

Example 4: VEST

- Extension of ITT thinking
- Encouragement is an “instrumental variable”

“Randomized Encouragement Design”

• Randomizing to encouragement to do something
  • Likely much lower adherence rates → much larger sample size needed
    • Adherence 40% vs. 10% instead of 90% vs. 3%?
  • But much easier??
• Works best if:
  • You are collecting outcome data passively (e.g., via EHR)
  • You don’t need consent before randomizing
  • The intervention could be considered “quality improvement”?

Example 4: VEST

Could we have used a Randomized Encouragement Trial to study the LifeVest?

(maybe?)
Example 4: VEST

Could we have used a Randomized Encouragement Trial to study the LifeVest?

(maybe?)

Passive data collection via EHR could make pragmatic RCTs much more feasible...
Learning Healthcare System

- A healthcare system that embeds pragmatic trials throughout delivery of healthcare, and learns as it goes...

- Learning Healthcare System vision
  - Every patient should be randomized! Multiple times!
  - Continuous learning and refinement of care
  - “A/B testing” embedded in EHR
    - Google does it, why shouldn’t we?
Learning Healthcare Systems

• UCSF CTSI is funding LHS demonstration projects

• Two projects funded:
  • Detection and alerts to avoid hypoglycemia events in the hospital
  • Using smart triggered standardized order sets in patients that may have COPD to improve quality of care + reduce readmissions and length of stay
Learning Healthcare Systems

• Steps for LHS projects
  • Identify an important health outcome that is easily measured by EHR query
  • Develop an intervention to improve it
  • Design a study to test whether it works
Learning Healthcare Systems

• Challenges for LHS projects
  
• Data work is HARD
  • Outcome assessment may be easy, but also need data to identify patients, create prediction algorithm, etc
  
• Intervention design is an art form
  • Alert fatigue
  • Workflow integration
  • User testing and involvement in design is critical
  
• Study design is fun
  • Unit of randomization?
  • Consent required or not*?
  • IRB approval pushing boundaries

Learning Healthcare Systems

• Even UCSF Health is talking this up
  • Learning Healthcare System is one of UCSF Health’s “True North Pillars” in a new strategic plan

• Institute for Computational Health Sciences, Atul Butte

• New cross-UC data initiative

http://www.ncbi.nlm.nih.gov/books/NBK53494/
Learning Healthcare Systems

• Hot topic nationally
  - IOM Reports from 2007 and 2012
    - See references below
  - PCORnet: a national network with a common data model → A national Learning Healthcare System?
    - PCORI has invested $100’s Millions
  - NIH CTSA Program is emphasizing EHR investments
  - NIH Collaboratory: Cutting edge, webinar series, living textbook

http://www.ncbi.nlm.nih.gov/books/NBK53494/
NIH Collaboratory https://www.nihcollaboratory.org/Pages/default.aspx
Science of EHR data

• All of this hinges on being able to use EHR data
Science of EHR data

• Some challenges in using EHR data

  • Collected for billing and clinical use, not research
    • Example: diagnosis data (clinicians know how bad this is)
  • Data structure designed to facilitate user interface
  • Every system is different
    • EPIC vs. Cerner vs. Others
    • EPIC installations are all customized
  • Data *collection* (not just the measurement value) inherently depends on health status
    • A recipe for selection bias
Science of EHR data

• Diagnosis data

  • Sources
    • Event-based (every visit, every lab)
    • On Problem List

  • Categories are sometimes hard
    • Garbage in vs. omitted (vs. time to get it right)

• ICD10 vs. ICD9
  • Different codes
  • Requires extreme detail
<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>ICD-10 Codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>26947</td>
<td>Diabetes mellitus arising in pregnancy</td>
<td>C24.419</td>
</tr>
<tr>
<td>17973</td>
<td>Diabetes mellitus associated with genetic syndrome</td>
<td>E11.69, Q99.9</td>
</tr>
<tr>
<td>17971</td>
<td>Diabetes mellitus associated with hormonal etiology</td>
<td>E11.69</td>
</tr>
<tr>
<td>50917</td>
<td>Diabetes mellitus associated with pancreatic disease</td>
<td>E11.68, K98.9</td>
</tr>
<tr>
<td>17974</td>
<td>Diabetes mellitus associated with receptor abnormality</td>
<td>E11.69</td>
</tr>
<tr>
<td>51794</td>
<td>Diabetes mellitus complicating pregnancy</td>
<td>O24.919, E11.9</td>
</tr>
<tr>
<td>52193</td>
<td>Diabetes mellitus complicating pregnancy, antepartum</td>
<td>C24.919</td>
</tr>
<tr>
<td>31543</td>
<td>Diabetes mellitus due to abnormal insulin</td>
<td>E13.9</td>
</tr>
<tr>
<td>57820</td>
<td>Diabetes mellitus due to cystic fibrosis</td>
<td>E08.9, E84.9</td>
</tr>
<tr>
<td>31542</td>
<td>Diabetes mellitus due to insulin receptor antibodies</td>
<td>E11.9</td>
</tr>
<tr>
<td>57926</td>
<td>Diabetes mellitus due to non-steroid drug</td>
<td>E09.9, T50.905A</td>
</tr>
<tr>
<td>57926</td>
<td>Diabetes mellitus due to non-steroid drug without complication</td>
<td>E09.9, T50.905A</td>
</tr>
<tr>
<td>1401294</td>
<td>Diabetes mellitus due to diabetic injury</td>
<td>E11.69, S36.209A</td>
</tr>
<tr>
<td>443814</td>
<td>Diabetes mellitus due to underlying condition</td>
<td>E08.9</td>
</tr>
<tr>
<td>442645</td>
<td>Diabetes mellitus due to underlying condition with diabetic amyotrophy</td>
<td>E08.44</td>
</tr>
<tr>
<td>442647</td>
<td>Diabetes mellitus due to underlying condition with diabetic cataract</td>
<td>E08.35</td>
</tr>
<tr>
<td>442649</td>
<td>Diabetes mellitus due to underlying condition with diabetic chronic kidney disease</td>
<td>E08.22</td>
</tr>
<tr>
<td>442651</td>
<td>Diabetes mellitus due to underlying condition with diabetic dermatitis</td>
<td>E08.620</td>
</tr>
<tr>
<td>442653</td>
<td>Diabetes mellitus due to underlying condition with diabetic mononeuropathy</td>
<td>E08.41</td>
</tr>
<tr>
<td>442655</td>
<td>Diabetes mellitus due to underlying condition with diabetic nephropathy</td>
<td>E08.21</td>
</tr>
<tr>
<td>442657</td>
<td>Diabetes mellitus due to underlying condition with diabetic nephropathy</td>
<td>E08.610</td>
</tr>
<tr>
<td>442659</td>
<td>Diabetes mellitus due to underlying condition with diabetic nephropathy</td>
<td>E08.610</td>
</tr>
<tr>
<td>442661</td>
<td>Diabetes mellitus due to underlying condition with diabetic peripheral angiopathy</td>
<td>E08.52</td>
</tr>
<tr>
<td>442663</td>
<td>Diabetes mellitus due to underlying condition with diabetic polyneuropathy</td>
<td>E08.42</td>
</tr>
<tr>
<td>442665</td>
<td>Diabetes mellitus due to underlying condition with foot ulcer</td>
<td>E08.62, L97.599</td>
</tr>
<tr>
<td>442666</td>
<td>Diabetes mellitus due to underlying condition with hyperglycemia</td>
<td>E08.65</td>
</tr>
<tr>
<td>442667</td>
<td>Diabetes mellitus due to underlying condition with hypoosmolarity with coma</td>
<td>E08.01</td>
</tr>
<tr>
<td>442668</td>
<td>Diabetes mellitus due to underlying condition with hypoosmolarity with nonketotic hyperglycemic-hypoosmolar coma</td>
<td>E08.00</td>
</tr>
<tr>
<td>442669</td>
<td>Diabetes mellitus due to underlying condition with hyperglycemia with coma</td>
<td>E08.641</td>
</tr>
<tr>
<td>442670</td>
<td>Diabetes mellitus due to underlying condition with hyperglycemia with coma</td>
<td>E08.649</td>
</tr>
<tr>
<td>442671</td>
<td>Diabetes mellitus due to underlying condition with ketoacidosis with coma</td>
<td>E08.11</td>
</tr>
<tr>
<td>442672</td>
<td>Diabetes mellitus due to underlying condition with ketoacidosis without coma</td>
<td>E08.10</td>
</tr>
<tr>
<td>442675</td>
<td>Diabetes mellitus due to underlying condition with mild nonproliferative diabetic retinopathy with macular edema</td>
<td>E08.321</td>
</tr>
<tr>
<td>442677</td>
<td>Diabetes mellitus due to underlying condition with mild nonproliferative diabetic retinopathy without macular edema</td>
<td>E08.329</td>
</tr>
<tr>
<td>442680</td>
<td>Diabetes mellitus due to underlying condition with moderate nonproliferative diabetic retinopathy with macular edema</td>
<td>E08.331</td>
</tr>
<tr>
<td>442682</td>
<td>Diabetes mellitus due to underlying condition with moderate nonproliferative diabetic retinopathy without macular edema</td>
<td>E08.339</td>
</tr>
<tr>
<td>442684</td>
<td>Diabetes mellitus due to underlying condition with periodontal disease</td>
<td>E08.630</td>
</tr>
<tr>
<td>442687</td>
<td>Diabetes mellitus due to underlying condition with proliferative diabetic retinopathy with macular edema</td>
<td>E08.351</td>
</tr>
<tr>
<td>442689</td>
<td>Diabetes mellitus due to underlying condition with proliferative diabetic retinopathy without macular edema</td>
<td>E08.359</td>
</tr>
<tr>
<td>442692</td>
<td>Diabetes mellitus due to underlying condition with severe nonproliferative diabetic retinopathy with macular edema</td>
<td>E08.341</td>
</tr>
<tr>
<td>442694</td>
<td>Diabetes mellitus due to underlying condition with severe nonproliferative diabetic retinopathy without macular edema</td>
<td>E08.349</td>
</tr>
<tr>
<td>442695</td>
<td>Diabetes mellitus due to underlying condition without complications</td>
<td>E08.9</td>
</tr>
<tr>
<td>442697</td>
<td>Diabetes mellitus due to underlying condition with diabetic autonomic neuropathy</td>
<td>E08.43</td>
</tr>
<tr>
<td>442699</td>
<td>Diabetes mellitus due to underlying condition with diabetic neuropathy</td>
<td>E08.40</td>
</tr>
<tr>
<td>442701</td>
<td>Diabetes mellitus due to underlying condition with diabetic retinopathy without macular edema</td>
<td>E08.319</td>
</tr>
<tr>
<td>442703</td>
<td>Diabetes mellitus due to underlying condition with diabetic retinopathy with macular edema</td>
<td>E08.314</td>
</tr>
</tbody>
</table>
Science of EHR data

• Medications data
  • Also coded, but it works better
  • But many sources
    • Prescribing event
    • Dispensing in facility
    • Dispensing at pharmacy
    • Med List

• How do you define what a patient is actually taking?
  • Med list (hard to find, not necessarily updated, single snapshot)
  • Start date easier than end date?
  • “Medication possession ratio”, etc
    • Requires pharmacy claims!
# Table 1

Methods of measuring adherence

<table>
<thead>
<tr>
<th>Methods</th>
<th>Data source</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indirect measurements used in research and administrative settings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MPR&lt;sup&gt;*&lt;/sup&gt;</td>
<td>Pharmacy claims</td>
<td>= (total days supplied)/(number of days between the first and last refills)</td>
</tr>
<tr>
<td>PDC&lt;sup&gt;*&lt;/sup&gt;</td>
<td>Pharmacy claims</td>
<td>= (total days supplied)/(number of days in refill interval)</td>
</tr>
<tr>
<td>Indirect measurements used in patient care settings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-report</td>
<td>Patient</td>
<td>Patient recalls medications taken in response to care team query</td>
</tr>
<tr>
<td>Questionnaire</td>
<td>Provider</td>
<td>Use of validated tool for adherence markers</td>
</tr>
<tr>
<td>Pill counting</td>
<td>Provider</td>
<td>Staff member reviews patient supply for doses remaining</td>
</tr>
<tr>
<td>Dose counting device</td>
<td>Device</td>
<td>Device includes electronic or manual counter that tracks doses released</td>
</tr>
<tr>
<td>Electronic-prescribing</td>
<td>PBM interface</td>
<td>Reports transmitted from a pharmacy benefit manager to provider usually via EMR link</td>
</tr>
<tr>
<td>Direct measurement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct observation</td>
<td>Provider</td>
<td>Patient receives and takes medication at health care facility</td>
</tr>
<tr>
<td>Drug levels and markers</td>
<td>Laboratory</td>
<td>Patient blood or urine sample tested</td>
</tr>
</tbody>
</table>

**Note:**

* Generally not used in direct patient care.

All abbreviations: PBM, pharmacy benefit manager; PDC, proportion of days covered; PBM, pharmacy benefit manager; EMR, electronic medical records.
Science of EHR data

• Laboratory data

  • Oldest and cleanest?
    • STOR data going back decades...

  • Units, normal ranges, comparability are often questionable
Slide from Mini-Sentinel and PCORnet
(courtesy of Lesley Curtis)

- 34 different result units for HbA1c
- 68 different result units for platelets
Science of EHR data

• Other numeric data...2 examples
  
  • Blood pressure
    • Context important?
    • In Emergency Dept vs. clinic, clinician-reported vs. dinamap, which arm, etc
    • Often need the “metadata”
  
  • Ejection fraction
    • Radiology/Cardiology reports – structured free text?
    • Code at UCSF – extracts from ~100 sources!
    • (thanks to Alvin Rajkomar)
Science of EHR data

- Hospitalization and death data
  - Diagnosis data coded more carefully?

- Useful...but not complete?
  - People use different hospitals
  - People die out of hospital
  - Closed system data better than open
    - Kaiser vs. UCSF
  - Claims data helps a LOT for this...but hard to get?
Science of EHR data

• Text data from written notes
  • Progress notes, discharge summary, nurses notes, etc
  • Extremely important information...but hard to “mine”!

• Whole science of natural language processing
  • Find concepts (e.g., heart attack, myocardial infarction, MI)
  • Identify “negation” language (e.g., “No ___”, “rule out”, “r/o”)
Science of EHR data

• Other special data types
  • EKG signals
  • Imaging files
Science of EHR data

• Computable phenotype concept

  • Coding of data can define useful research measurements
    • Diabetes (diagnosis + meds + FASTING blood sugar level?)
    • Hypertension (uncontrolled, treated, resistant, etc)
    • Hospitalization for myocardial infarction (Diagnosis + EKG + NLP?)

• NEVER perfect
  • Validation by chart review, etc...
  • Variations in CP’s will provide different tradeoffs in terms of sensitivity and specificity (depends on research study needs)
Science of EHR data

• Data Models
  • Take source data and modify
    • Simplified data model (combining sources, etc), easier to use
    • **Common data model** for use across institutions
      • Allows multi-site studies
      • Allows crowd-sourced computable phenotype development
  
• “Extract-Transform-Load” - ETL
  • Requires data manipulation, joining, concatenating datasets
  • Decisions!
  • Information loss at every step
Figure B1. Clinical Data Flow, Transformation, and Usage in our EPIC-based EHR, with planned enhancements.
Science of EHR data

• Some important UCSF-supported data models to know about:

  • Caboodle (AKA Cogito, Star, “Constellation”)
    • EPIC product, simplification of APEX data
    • UCSF has de-identified flat files for many Caboodle tables!

  • OMOP (hosted by OHDSI)
    • Open source community effort
    • Competes with i2b2 https://www.i2b2.org/about/intro.html

  • PCORnet Common Data Model
    • National network funded to do EHR work
      http://www.pcornet.org/pcornet-common-data-model/
Real-time data needs (e.g., ED recruiting)

Clinical/Ops “dashboards” and research exploration tools

Research datasets

IBM Infosphere Metadata, reusable code library, etc

Figure B1. Clinical Data Flow, Transformation, and Usage in our EPIC-based EHR, with planned enhancements
Science of EHR data

• “Big data” approaches to EHR data
  • Define computable phenotype and use prediction tools
    • Neural networks, etc
    • Let’s just sic Google on it...?

• Unsupervised – will it work?
  • Do we use Clarity 10,000 tables or Caboodle 200 or OMOP 20?
  • NLP?
  • Maybe if we choose the right problem?
Science of EHR data

• Accessing UCSF APEX data

• Consultations from CTSI
  • https://accelerate.ucsf.edu/consult

• Going directly to UCSF IT (AKA Academic Research Systems = ARS)
  • https://it.ucsf.edu/services/clinical-data-research-consultations-formerly-threds-and-idr-data-extraction-service
Take Home Points

• Potential win-win of large, simple, pragmatic trials
  • Cheaper, and more relevant for decision-making

• Learning healthcare system concept
  • Embedding trials in healthcare delivery
  • UCSF efforts

• Using electronic health record data
  • Critical for LHS vision, but hard!
  • Active area of development at UCSF
  • Resources available; some now and more coming soon
Take Home Points

• Potential win-win of large, simple, pragmatic trials
  • Cheaper, and more relevant for decision-making

• Learning healthcare system concept
  • Embedding trials in healthcare delivery
  • UCSF efforts

• Using electronic health record data
  • Critical for LHS vision, but hard!
  • Active area of development at UCSF
  • Resources available; some now and more coming soon

Questions?
Extra slides to follow
Learning Healthcare Systems

• Quality Improvement Research
  • QI vs. Research: OVERLAP, not spectrum/dichotomy
  • Too many projects try to avoid IRB approval
    • Better to use good research methods (e.g., randomization)
    • Waiver/modification of consent is often justifiable to IRBs

• Randomized Quality Improvement Trial*
  • A special beast
  • Equipoise not relevant?
  • Encouragement design especially reasonable?