Real-world data in pragmatic trials

Harold C. Sox, MD
The Patient-Centered Outcomes Research Institute
Washington, DC

Presenter Disclosure Information

In compliance with the accrediting board policies, the American Diabetes Association requires the following disclosure to the participants:

Name of Presenter
Harold C. Sox, MD

Disclosed no conflict of interest.

The content of the talk does not necessarily reflect the policies of PCORI or its Board of Governors.
Plan of the talk

1. The concepts and conduct of a pragmatic trial.
2. Real-world data: definitions and sources
4. Use of real-world data in ADAPTABLE: advantages and pitfalls.
5. Challenges to interpreting pragmatic clinical trial results.

Pragmatic Trials: history

• 1967: Schwartz and Lellouch (J Chronic Disease): pragmatic and explanatory attitudes toward trials.
  • “This article makes no pretention to originality, nor to the provision of solutions; we hope we have clarified certain issues to the extent of encouraging further discussion.”
• 1983-85: GISSI trial of thrombolysis in acute MI.
• 2005: PRECIS-1
  • “Finally, we stress that this article, building on earlier work from multiple investigators, describes a “work in progress.”
• 2013: PRECIS-2
<table>
<thead>
<tr>
<th>Pragmatic</th>
<th>Efficacy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broad eligibility</td>
<td>Narrow eligibility</td>
</tr>
<tr>
<td>Flexible interventions</td>
<td>Adhere to instructions</td>
</tr>
<tr>
<td>Typical practitioners</td>
<td>Expert practitioners</td>
</tr>
<tr>
<td>Follow-up visits only as needed for care</td>
<td>Frequent follow-up visits</td>
</tr>
<tr>
<td>Objective clinical outcome</td>
<td>Surrogate outcomes</td>
</tr>
<tr>
<td>Usual efforts to insure adherence</td>
<td>Close monitoring for adherence</td>
</tr>
<tr>
<td>Addresses a clinical decision</td>
<td>Addresses mechanism and concept</td>
</tr>
<tr>
<td>Per Protocol analysis</td>
<td>Intention to treat analysis</td>
</tr>
</tbody>
</table>

Figure 1: The blank “wheel” of the pragmatic–explanatory continuum indicator summary (PRECIS) tool. “E” represents the “explanatory” end of the pragmatic–explanatory continuum.

Kevin E. Thorpe et al. CMAJ 2009;180:E47-E57

©2009 by Canadian Medical Association
Steps in pragmatic study design

- **Step 1**: What design approach are you taking?
  - Can this intervention work under ideal conditions?
  - Does it work under usual conditions of care?

- **Step 2**: Consider your design choices for each of the 8 PRECIS domains
  - Imagine what each domain would look like for a very pragmatic design. A very explanatory design. Choose.

- **Step 3**: Use the PRECIS wheel to map your study
  - Rate your choice for each domain (1-5 scale)

- **Step 4**: Use the study map to revisit your choices.

Real-world data in clinical trials

Definitions

• **Real-world data:**
  • Data obtained outside the context of RCTs generated during routine clinical practice. (ISPOR-ISPE Special Task Force on Real-World Evidence in Health Care Decision Making, 2107).

• **Pragmatic trial:**
  • A trial done to determine the effects of an intervention under the circumstances in which it will be applied in practice.

• **Explanatory trial:**
  • A trial done to determine the effects of an intervention under ideal circumstances.
SOURCES OF REAL-WORLD DATA

- Electronic Health Records (EHR)
- Claims and billing data
- Product and disease registries
- Healthcare applications on personal digital devices

Source: Matthew Roe. Clinical Leader, August 2017

ADAPTABLE: THE ASPIRIN DOSE QUESTION

- Patients with established coronary heart disease
- Compares aspirin at two doses: 81 mg per day and 325 mg per day
- Conducted through a Clinical Data Research Network (CDRN) that is part of the PCORnet “network of networks.”
  - Collaboration of 8 CDRNS representing 30 health care systems
- **Common Data Model**: a library of descriptors (and their definitions) taken from the study sites. **Purpose**: to translate diverse definitions of an item into one representation in the study data base
ADAPTABLE: STUDY CONDUCT

• **Identifying participants**: search EHRs or administrative claims database for eligible patients ("computable phenotype")

• **Outreach**:
  - **Patients**: email, patient portals, alerts to providers, tablet-based recruiting at patient visits.
  - **Providers**: inform; opt-out; pop-up to remind provider during a visit with a potentially eligible patient

• **Enrollment**: Interested patients get a personal code to access the ADAPTABLE web portal for confirmatory screening, e-consent, self-randomization

• **Follow-up**: Enrollees reminded to return to the portal to report study endpoints

Validation of recruitment methods

• Using the computable phenotype on health systems data to locate potential enrollees was reliable.
  - The authors reviewed the medical records of 185 patients identified by the computable phenotype as study-eligible.
  - 85% of those identified as study eligible by the computer search were confirmed as eligible by medical record review.
  - Probability of being eligible if computable phenotype is positive = 0.85.

E Fishman et al: poster presented at PCORI Annual Meeting. October 2018
ADAPTABLE: outcomes

- Outcomes: hospitalization for MI or stroke or death
- Each PCORnet partner network maps outcome definitions to a common format (THE COMMON DATA MODEL) on a platform that searches participating hospitals’ EHR.
- Surveillance of claims data obtained from health plan networks or Medicare/Medicaid.
- Death from Social Security death index.
- Study personnel contact participants who miss their e-follow-up dates

Sources of inaccurate real-world data

- Missing data
  - Providers do not consistently get a core condition-specific data set on every patient with suspected of the condition.
  - The patient gets care outside their usual health system.
- Inaccurate transcription by a multi-tasking health professional
- Terminal digit bias (tends to give even numbers)
- Rounding error (144/70 → 140/70)
- Extreme values inconsistent with prior ones
Validation of outcome ascertainment methods

- Patients seek care beyond the recruitment network.
- Relying on data from the network’s associated health systems (EHR, insurance claims) will miss 30-40% of outcomes.
- Using insurance claims from inside and outside the research network is more reliable.


Dealing with problems in the design and analysis of (pragmatic) randomized trials
In principle, random assignment to a treatment guarantees that it is the cause of any differences between study arms.

Missing data: a major problem for interpreting study results

• Consider a RCT comparing a drug with intensive physical therapy and physical exercise for chronic low back pain.
  • 100 patients per study arm
  • 6 months of therapy
  • The main outcome is a questionnaire about back pain and its effects.
  • At the end of the study all 100 patients in the drug arm—but only the 60 patients who completed the PT/exercise intervention—are available for follow-up.
  • The back pain scores are better in the drug arm.

• What do the results say about the comparative effectiveness of the two interventions in the study population?
Missing data: a major problem for interpreting study results

- Consider a RCT comparing a drug with intensive physical therapy and physical exercise for chronic low back pain.
  - 100 patients per study arm
  - 6 months of therapy
  - The main outcome is a questionnaire about back pain and its effects.
  - At the end of the study all 100 patients in the drug arm—but only the 60 patients who completed the PT/exercise intervention—are available for follow-up.
  - The back pain scores are better in the drug arm.

- What do the results say about the comparative effectiveness of the two interventions in the study population?
  - Breaking randomization → lose the guarantee that unmeasured confounders will be balanced between study arms.

<table>
<thead>
<tr>
<th>Outcome measure</th>
<th>Ideal scenario</th>
<th>Real-life scenario</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcome is a measure of quality of life obtained by survey or interview</td>
<td>The patient responds the first time to a phone call, is fluent in English, and cooperative</td>
<td>The patient doesn’t have a phone, is not fluent in English, and is does not keep medical appointments.</td>
</tr>
<tr>
<td>Outcome is a laboratory test result obtained from the EHR</td>
<td>The patient goes to a physician regularly, the physician remembers to order the test, the test gets done, and the result gets into the EHR</td>
<td>The patient is too frail to go to the hospital to get tested, has no money to pay for the test, has no accessible veins, and the doctor ignores the pop-up reminder to order the test.</td>
</tr>
</tbody>
</table>
Approach to missing data

• “the assumption that analysis methods can compensate for such missing data are not justified, so aspects of trial design that limit the likelihood of missing data should be an important objective.” (R Little, NEJM)

• Bottom line:
  • Try very, very hard to contact all patients
  • Soft-pedal study results when the drop-out rate exceeds 25%.


Approach to missing data: imputation

• An analytical approach: use study data that are available to impute the missing values.
  • This assumes missingness at random (MAR): that observed data can account for differences in the distribution of variables in observed and missing cases.
  • MAR means that the result on the missing item can be inferred from the same item on study participants who did not drop out and were similar at baseline to those with the missing item.
  • The MAR assumption can be verified for observed variables but not for unmeasured variables.
  • A solution is to posit an unmeasured variable, assume various values for it, and see if the study results change (sensitivity analysis).
Per-Protocol Analyses

• Patients want to know the effect of the intervention when taken as directed (i.e., in patients who adhere to the study protocol).

• Post-randomization events (e.g., stopping the study medication) can lead participants to depart from the study protocol.
  • Lost to follow-up → use available outcome data
  • Poor adherence to the intervention → adjust outcome measures for prognostic factors that influence adherence (so that outcomes reflect perfect adherence).

• To adjust for post-randomization prognostic factors, the researcher must track them during the study.


Intent to treat vs. complete case (per protocol) analysis

• Intent to treat:
  • Analyze the data on all randomized patients.
  • Use outcome data from the last contact before lost to follow-up. LCOF
  • Results tell you the average result over study completers and study drop-outs.
  • Drop-outs may differ from completers in ways that affect study outcomes.
  • Results can mislead patients who make treatment decisions.

• Per protocol analysis
  • Analyze the data so that it reflects perfect protocol adherence.
  • Must adjust study outcomes for post-randomization prognostic factors that affect adherence (g-methods) → study arms have equivalent adherence.
  • Results tell you what to expect if a patient adheres to the treatment.

Plan of the talk

1. The concepts and conduct of a pragmatic trial.
2. Real-world data: definitions and sources
4. Use of real-world data in ADAPTABLE: advantages and pitfalls.
5. Challenges to interpreting clinical trial results.

Summary

• We have discussed the use of real-world data in pragmatic trials.
• Pragmatic trials and efficacy trials are at the opposite ends of a continuous spectrum.
• ADAPTABLE is a very large network-based pragmatic RCT that uses outcome data collected as part of routine care.
• Missing data and poor adherence are big problems in clinical research.
  • The best approach to missing data is prevention. Analytic retro-fixes depend on unverifiable assumptions.
  • Adjusting for prognostic factors that affect adherence is a possible answer to non-adherence.