

Perspective

Design and analytic considerations for using patient-reported health data in pragmatic clinical trials: report from an NIH Collaboratory roundtable

Frank W. Rockhold ^{1,2}, Jessica D. Tenenbaum,^{1,2} Rachel Richesson,^{1,3} Keith A. Marsolo,⁴ and Emily C. O'Brien^{1,5}

¹Duke Clinical Research Institute, Durham, North Carolina, USA, ²Department of Biostatistics and Bioinformatics, Duke University School of Medicine, Durham, North Carolina, USA, ³Duke University School of Nursing, Durham, North Carolina, USA, ⁴Duke University School of Medicine, Durham, North Carolina, USA, and ⁵Population Health Sciences, Duke University School of Medicine, Durham, North Carolina, USA

Corresponding Author: Emily C. O'Brien, PhD, Duke Clinical Research Institute, PO Box 17969, Durham, NC 27715, USA (emily.obrien@duke.edu)

Received 5 July 2019; Revised 10 October 2019; Editorial Decision 15 December 2019; Accepted 20 December 2019

ABSTRACT

Pragmatic clinical trials often entail the use of electronic health record (EHR) and claims data, but bias and quality issues associated with these data can limit their fitness for research purposes particularly for study end points. Patient-reported health (PRH) data can be used to confirm or supplement EHR and claims data in pragmatic trials, but these data can bring their own biases. Moreover, PRH data can complicate analyses if they are discordant with other sources. Using experience in the design and conduct of multi-site pragmatic trials, we itemize the strengths and limitations of PRH data and identify situational criteria for determining when PRH data are appropriate or ideal to fill gaps in the evidence collected from EHRs. To provide guidance for the scientific rationale and appropriate use of patient-reported data in pragmatic clinical trials, we describe approaches for ascertaining and classifying study end points and addressing issues of incomplete data, data alignment, and concordance. We conclude by identifying areas that require more research.

Key words: electronic health records, patient reported outcome measures, pragmatic clinical trials as topic, randomized controlled trials as topic, patient-generated health data

INTRODUCTION

Researchers are increasingly using data from electronic health records (EHRs) in observational and interventional research, including pragmatic clinical trials, to improve the real-world relevance, speed, and efficiency of clinical research. In this article, we refer to pragmatic clinical trials as interventional studies embedded within health care systems and designed to provide evidence for the effectiveness of interventions in real-world settings (Table 1). In addition, we refer to EHRs as a source of health care data for research, and many of the issues we discuss also apply to claims data. This approach to randomized trials has been reviewed in detail

elsewhere.^{1–4} The Food and Drug Administration's commitment to "real-world evidence" for regulating drug safety includes the use of pragmatic trials.⁵ GlaxoSmithKline's Salford Lung Study⁶ was the first Phase 3 pragmatic clinical trial used to support marketing of a new drug.

Pragmatic clinical trials often leverage data from EHRs to facilitate recruitment of appropriate patients, characterize patients for analysis, and assess study end points. The use of EHRs for these purposes improves the feasibility and scalability of pragmatic clinical trials. Yet, EHR data can introduce challenges with completeness, accuracy, and timeliness of data captured during routine clinical

Table 1. Potential differences in key design features of pragmatic and traditional randomized controlled trials

Traditional trial		Pragmatic trial
Intentionally homogeneous to maximize treatment effect	Eligibility criteria	Heterogeneous—representative of normal treatment population
Randomization and blinding	Randomization and blinding	Randomization and rarely blinding
Clinical measures, intermediate end points, composite end points, clinical outcomes	End points	Clinical outcomes, patient-reported outcomes, quality of life, resource use
Protocol defines the level and timing of testing. Physicians blinded to data	Tests and diagnostics	Measured according to standard practice
Fixed standard of care of placebo	Comparison intervention	Standard clinical practice
Conducted only by investigators with proven track record	Practitioner expertise	Employment of a variety of practitioners with differing expertise and experience
Visit schedule and treatment pathway defined in the protocol	Follow-up	Visits at the discretion of physician and patient
Patients wishing to change treatment must withdraw from the study	Continuity	Standard clinical practice—switching therapy according to patient needs
Compliance is monitored closely	Participant compliance	Passive monitoring of patient compliance
Close monitoring of adherence	Adherence to study protocol	Passive monitoring of practitioner adherence
Intent to treat, per protocol and completers	Analysis	All patients included

care.^{7,8} Fragmented care can influence the performance of phenotype algorithms, as previously demonstrated, for example, in diabetes detection.⁹ Furthermore, data in the EHR may lack the level of specificity required for a pragmatic trial. Outcomes must be defined clearly, measured objectively, and collected completely, because they are used to determine whether an intervention is effective.

Patient-reported health (PRH) data can enhance pragmatic trials by augmenting EHR data for study end points. We define PRH data to include patient- or caregiver-reported information that may also exist in the EHR, including hospitalizations, comorbid conditions, and medications. For our purposes, PRH data do not include subjective patient-reported outcomes, such as symptoms. PRH data can be used to confirm or supplement EHR data to address data fragmentation arising from the use of data from multiple provider organizations. However, if PRH data contradict EHR data, analysts must make decisions about which data to accept. In many cases, no single source offers a clear advantage, particularly because clinical events may happen outside the health system and PRH data are vulnerable to recall bias.

Despite the importance of PRH data for pragmatic research, best practices for integrating PRH data with clinical data sources have not been defined. There is some guidance for analyzing PRH data,¹⁰⁻¹³ but few road maps exist to guide researchers through questions of when, where, and under what circumstances to integrate PRH data for real-world evidence generation.

The National Institutes of Health (NIH) Health Care Systems Research Collaboratory organized a roundtable discussion to address best practices for capturing PRH data in pragmatic studies and optimal analytic approaches for integrating PRH with other data sources. The group of statisticians and informaticists met in person in September 2017 to develop the guidance. As a result of these discussions, we produced consensus findings focused on analysis and integration of PRH data in clinical trial analyses.

KNOWING WHEN PRH DATA MAY BE NEEDED IN A PRAGMATIC TRIAL

Designers of pragmatic trials often must consider alternative data sources, including patient reports, due to the incompleteness of EHR data for research purposes. Decisions made during the trial design

phase can improve the validity of data capture and the robustness of the data used for analysis. We present several key considerations and offer solutions to achieving concordance and approaches to minimize misalignment between data sources.

Approaches for ascertaining and classifying end points in pragmatic trials

There are unique features related to the ascertainment of outcomes in randomized pragmatic clinical trials.^{2,14} A key consideration is whether randomization introduces differential misclassification and potential ascertainment bias in trials that are open-label. Another key feature is the nature of the control, that is, whether it is an “active” controlled trial in which 1 of the therapies is known to be effective in the trial population or a treatment compared to “no treatment,” both being given in addition to standard therapies.

The sensitivity and specificity of outcome ascertainment are important and depend on trial design. Even under optimal design conditions, pragmatic clinical trials typically have lower specificity (true negative rate) and sensitivity (true positive rate) with respect to event capture than traditional randomized clinical trials. Depth and auditability also are not the same as in a standard clinical trial with a well-designed, stand-alone case report form. The lower expected outcome sensitivity in cardiovascular outcome trials such as ADAPTABLE¹ may be addressed by using other information in the EHR and by mimicking the adjudication process of a randomized controlled trial. Specificity is more difficult to ensure and often requires supplemental resources, such as claims data, National Death Index data, and patient-reported information to ensure events are not missed.

GAPS IN DATA: EVIDENCE OF ABSENCE OR ABSENCE OF EVIDENCE?

The choice to use PRH data to supplement EHR data in a pragmatic clinical trial depends on the nature of expected information gaps in the EHR data. Here we describe several scenarios in which outcome ascertainment may be incomplete because of limitations in EHR data capture (Table 2). We also offer several potential approaches to resolution for each limitation type.

Table 2. Information gaps and approaches to resolution in patient-reported health data

Cause	Example	Approach to resolution
Event occurred at different health system	Patient hospitalized while on vacation, heart attack captured in EHR of external health system	Include questions on recent hospitalizations in follow-up contacts directly with patient via call center or web portal
Event occurred, but not recorded	Patient has racing pulse, but does not seek treatment; patient in inpatient unit with vitals recorded every 15 minutes. Racing pulse occurs and subsides in between recordings.	Consider use of PRH data as the primary data source for events that are unlikely to be routinely available in EHR data
Event occurred, does not appear in data source due to other events of higher priority	Event recorded as diagnosis, but does not appear on bill because of other diagnoses with higher reimbursement	Sensitivity analysis treating PRH data as the primary source of information on events for scenarios where patient reports event not apparent in billing data. Consider concordance with EHR data for final determination
Event occurred, but recorded as different/unexpected code or field	1. Procedure coded as X when trial is looking for code Y 2. Vital signs typically recorded in standard flowsheet, but clinic has custom field to use for their own patients 3. EHR is upgraded and new fields created—trial is unaware.	Sensitivity analysis treating PRH data as the primary source of information on events for scenarios where patient report contradicts EHR data
Event not recorded reliably	1. Trial is looking for height at most recent encounter; patient was seen the previous week & had height recorded then. Height not recorded a second time. 2. Trial is looking at smoking status, but field is not recorded reliably for young adults seen in pediatric clinics	Consider targeted capture of patient-reported data in scenarios where key data elements are at high risk of missingness; look for data collected previously using a look-back period as defined in the study protocol (ie, height within 2 weeks of visit)
Lag in process to extract EHR data to research database	Process to refresh research database runs quarterly; database only includes EHR data at least a month old.	Employ more frequent, consistent PRH data capture for trial monitoring purposes with confirmation of safety signals through targeted medical record review

Data collected sometimes, not always

A given value, such as vital signs, is measured in 3 main cases: during an outpatient visit to a doctor's office, as recurring values as inpatients, or as a specific cause for concern. Effectively, vital signs are "missing" from a patient's record for most of his or her life. Therefore, "missingness" (or, conversely, availability) of values is likely to differ by health status.

Data not collected reliably

A given variable may be collected as part of the standard of care for a chronic disease (for example, a physician's global assessment of disease status for patients with inflammatory bowel disease). Since this variable is available as part of the standard of care, it is expected to be present at every visit. However, because of time pressures or other issues, this field may not be captured consistently.

Data collected elsewhere

Patients may move between cities or states, or seek medical care while traveling. Even those who have lived in the same place their entire life are likely to see more than 1 health care provider. Therefore, data regarding encounters, treatments, or even major procedures or hospitalizations may exist, but in a different EHR that is not accessible to the researcher. For example, a patient who grew up in California but now receives care in New York as an adult may report having had the chicken pox as a child. In this case, the patient's self-report of chicken pox is likely to be more reliable than her adult provider's system, which is unlikely to have evidence of her infection in childhood.

Collected but not easily findable

Some data elements may be found only in free text or may be structured data found in fields other than the 1 expected.⁸ For example, a

researcher might be interested in ejection fraction information. However, a numeric value for ejection fraction may only be captured in text notes at some centers, in a multitude of different ways of indicating its meaning. Even if data are collected, they may be stored in different places in the record. For example, a drug allergy may be listed as an allergy, a diagnosis, or an item on a problem list. Finally, even structured data representing a single concept may have different naming conventions, depending on the system in which the data were recorded. Standard terminologies, such as SNOMED and LOINC, provide a common language to enable identification of the same clinical concepts across disparate systems.

Data not collected

Some data elements that might be useful are not collected at all, or at least not recorded. These often include environmental exposures, lifestyle information beyond basic drug, smoking, and alcohol use, and socioeconomic factors.¹⁵ In addition, use of over-the-counter medications may not be recorded in the medical record.¹⁶ These items may not be considered "missing," because they were never intended to be there in the first place, useful though they would be to the researcher.

Latency

Data latency is an important issue because, while information is uploaded quickly to billing systems from EHRs, it may be less current than information collected from other sources due to the additional step of transformation into a research-ready format. This introduces special considerations for ongoing surveillance of event rates in a distributed research network and may affect the functioning of the data monitoring committee.

DATA CONCORDANCE

Even when data elements can be aligned conceptually and have values present from both sources, there still can be discordance in values.¹⁷ The added utility of patient-reported data for capture of health events depends on the condition being studied and the characteristics of the patient population. The optimal approach for comparison of patient-reported data and other data sources varies across studies.^{11–13} As a result, there is insufficient evidence to support concrete recommendations for the best approach for using PRH data to classify event status in specific populations. More work is needed to identify which populations and study questions are most vulnerable to recall bias and misclassification to inform data collection strategies for pragmatic clinical trials.

A useful consideration in understanding reliability is examining the original source of the data in each case. Depending on the size of the study, the degree of discordance, and the resources available, a decision must be made for each data element regarding whether to choose 1 source as “truth” for all patients, or whether to adjudicate on an individual basis. If it is feasible to consider adjudication on a patient-by-patient basis, values or other variables may support use of the other source or 1 source or the other. For example, a person’s age or sex may rule out positive pregnancy status. These efforts are likely to alter the level of pragmatism and efficiency in a given study.

Though some types of data will require choosing 1 source over another, in other cases the sources may be combined. A combination of EHR data and the patient’s memory may represent the best triangulation of truth for questions regarding whether the patient has ever had a given diagnosis or procedure.

CASE STUDY: APPROACHES FOR END POINT ASCERTAINMENT AND CONFIRMATION IN ADAPTABLE

Details of the ADAPTABLE study design have been published elsewhere.¹⁴ The objective of the trial is to compare the effectiveness of low-dose vs high-dose aspirin on death, nonfatal myocardial infarction, nonfatal stroke, and major bleeding. It is randomized at the patient level, and outcomes are based primarily on patients’ medical records augmented by an online patient portal, where patients report events, and a call-in system for longitudinal follow-up to confirm clinical endpoints.

For the final analysis in ADAPTABLE, the protocol prespecifies that EHR will be used as the final adjudicator of presence or absence of events for study end points (with the exception of death), as a priori the study investigators believed this would be the most complete record at the beginning of the study. The validity of this assumption will be evaluated when the final data are analyzed.

EVIDENCE GAPS/FUTURE DIRECTIONS

There are several key evidence gaps which, if addressed, have the potential to greatly enhance understanding of how and where PRH data should be used in pragmatic clinical trials. As noted above, these include data collected sometimes but not always, data needed for research but not collected in routine care, data latency, and data collected outside the patient’s primary health system. While many have attempted to quantify the incomplete capture of all hospitalized events of interest for a given patient when using EHR data, little data are available quantifying the degree of incompleteness due to patients receiving acute care outside of a given system. In ADAPT-

ABLE, researchers are exploring this and plan to report results, including how data incompleteness varies according to the event of interest (eg, myocardial infarction vs major bleeding), location, population-level demographic characteristics, and availability of other nearby treatment centers.

CONCLUSION

PRH data should be considered as part of a larger set of information collected to enhance overall event capture and classification in pragmatic clinical trials. In this set of complementary data sources, PRH data may be integrated as part of a hierarchical approach to event classification. In some scenarios, patient-reported events may be followed with explicit confirmation via hospital bill collection or corroboration with a treating physician. In others, PRH data may be the only source of information available about health events for a given period of follow-up. There is little empirical evidence for how analytic decisions on how to treat patient-reported events that are in conflict with event data from other sources affects overall event capture with respect to sensitivity, specificity, and overall accuracy. Future studies evaluating the performance of event classification algorithms with integration of PRH data under commonly observed scenarios in pragmatic trials are needed.

FUNDING

Research reported in this publication was supported by the National Center for Complementary and Integrative Health of the National Institutes of Health under award number U54AT007748. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

AUTHOR CONTRIBUTIONS

FWR, KM, and ECO conceived of the work. All authors contributed to the drafting of the manuscript, made critical revisions for important intellectual content, and provided final approval of the version to be published.

CONFLICT OF INTEREST STATEMENT

None declared.

REFERENCES

1. Ford I, Norrie J. Pragmatic trials. *N Engl J Med* 2016; 375 (5): 454–63.
2. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutic trials. *J Chronic Dis* 1967; 20 (8): 637–48.
3. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ* 2015; 350: h2147.
4. Zwarenstein M, Treweek S, Gagnier JJ, *et al*. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ* 2008; 337: a2390.
5. 21st Century Cures Act. Pub L 114-255, 130 Stat 1033–1344.
6. New JP, Bakerly ND, Leather D, Woodcock A. Obtaining real-world evidence: the Salford Lung Study. *Thorax* 2014; 69 (12): 1152–4.
7. Richesson RL, Horvath MM, Rusincovitch SA. Clinical research informatics and electronic health record data. *Yearb Med Inform* 2014; 9: 215–23.

8. Hersh WR, Weiner MG, Embi PJ, *et al.* Caveats for the use of operational electronic health record data in comparative effectiveness research. *Med Care* 2013; 51: S30–7.
9. Wei WQ, Leibson CL, Ransom JE, *et al.* Impact of data fragmentation across healthcare centers on the accuracy of a high-throughput clinical phenotyping algorithm for specifying subjects with type 2 diabetes mellitus. *J Am Med Inform Assoc* 2012; 19 (2): 219–24.
10. Yasaitis LC, Berkman LF, Chandra A. Comparison of self-reported and Medicare claims-identified acute myocardial infarction. *Circulation* 2015; 131 (17): 1477–85; discussion 85.
11. Eichler GS, Cochin E, Han J, *et al.* Exploring concordance of patient-reported information on PatientsLikeMe and medical claims data at the patient level. *J Med Internet Res* 2016; 18 (5): e110.
12. De-Loyde KJ, Harrison JD, Durcinoska I, *et al.* Which information source is best? Concordance between patient report, clinician report and medical records of patient co-morbidity and adjuvant therapy health information. *J Eval Clin Pract* 2015; 21 (2): 339–46.
13. Basch E, Bennett A, Pietanza MC. Use of patient-reported outcomes to improve the predictive accuracy of clinician-reported adverse events. *J Natl Cancer Inst* 2011; 103 (24): 1808–10.
14. Johnston A, Jones WS, Hernandez AF. The ADAPTABLE trial and aspirin dosing in secondary prevention for patients with coronary artery disease. *Curr Cardiol Rep* 2016; 18 (8): 81.
15. Miranda ML, Ferranti J, Strauss B, *et al.* Geographic health information systems: a platform to support the triple aim. *Health Aff (Millwood)* 2013; 32 (9): 1608–15.
16. Schmiemann G, Bahr M, Gurjanov A, *et al.* Differences between patient medication records held by general practitioners and the drugs actually consumed by the patients. *Int J Clin Pharmacol Ther* 2012; 50 (08): 614–7.
17. Tisnado DM, Adams JL, Liu H, *et al.* What is the concordance between the medical record and patient self-report as data sources for ambulatory care? *Med Care* 2006; 44 (2): 132–40.