Patient-Centered Outcomes Research Institute Proposed New and Revised Methodology Standards (2016)

AcademyHealth is pleased to submit comments to the Patient-Centered Outcomes Research Institute's (PCORI's) proposed new and revised Methodology Standards, updated from the first set of standards initially released in December 2012.

AcademyHealth thanks the PCORI Methodology Committee for reevaluating these standards, understanding that as the field evolves, so, too, must the standards under which patient-centered outcomes researchers operate. The following represent the newly proposed Standards, with the comments AcademyHealth submitted on the PCORI website underneath.

1. Standards for Formulating Research Questions

What Changed:

Revisions consist of minor changes in wording to help define certain terms and to streamline and clarify the standards. The content of the standards remains essentially the same.

RQ-1: Identify gaps in evidence

Analysis of gaps in the evidence based on systematic reviews should be used to support the need for a proposed study. If a systematic review is not available, a systematic review should be performed using accepted standards in the field (see standard SR-1), or a strong rationale should be presented for proceeding without a systematic review. In the case where a systematic review is not possible, the methods used to review the literature should be explained and justified.

As proposed, the call to use 'analysis of gaps in the evidence based on systematic reviews' is much more complex than the requirement in the prior version of the Standard. Given the increased emphasis on using systematic reviews to identify evidence gaps as well as to ensure that PCORI's documents are as useful and accessible as possible, PCORI should consider making available on its website the best sources of systematic reviews, such as Cochrane, formally known as the Cochrane Collaboration (http://www.cochrane.org/) and Canada's Health Systems Evidence (https://www.healthsystemsevidence.org/), so researchers have a more solid foundation from which to begin their study.

Furthermore, when conducting systematic reviews, the central issue is not necessarily about what is possible, but what is useful and worthwhile. Therefore, we propose that the last sentence be: "In cases where incremental systematic reviews may not be useful over existing literature or in which the effort needed is not worthwhile, the methods used to review the literature should be explained and justified."

RQ-2: Develop a formal study protocol

Studies should include a formal protocol specifying at least one purpose for which
the data were collected (e.g., effectiveness, safety, natural history of disease, quality improvement); data sources and linkage plans, if any; data feasibility and quality,
measure(s) of effect; and use of any standardized data dictionaries (nationally or internationally accepted).

RQ-3: Identify specific populations and health decision(s) affected by the research

To produce information that is meaningful and useful to people when making specific health decisions, research proposals and protocols should describe:
the specific health decision the research is intended to inform;
the specific population for whom the health decision is pertinent; and
how study results will inform the health decision.

The PICOTS typology to which PCORI refers later in its Methodology Standards is used in systematic reviews for precisely this purpose of identifying and characterizing evidence gaps from systematic reviews. Therefore, PCORI should cite the framework in this Standard and advise researchers to cite it when producing the requested information. Additionally, AcademyHealth recommends changing the word 'population' in the second bullet to 'populations' (plural) both to maintain consistency with the Standard's title and to encompass as wide a group as necessary for a given study.

RQ-4: Identify and assess participant subgroups

In designing studies, researchers should identify participant subgroups and explain why they are of interest, preferably based on prior data. Where feasible, the study should have adequate precision and power to reach conclusions specific to these subgroups.

Within this Standard, AcademyHealth felt it was unclear whether PCORI was articulating whether researchers should not do subgroup analysis if there was not adequate precision and power to reach conclusions or whether that information should simply be reported. Clarification from PCORI on this point would be helpful. The clarification could distinguish when subgroup analyses are aimed at providing definitive advice for the effectiveness of particular interventions or when such analyses may be useful for learning more about subgroup issues and generating hypotheses for future research.

RQ-5: Select appropriate interventions and comparators

The interventions and comparators should correspond to the actual healthcare options for patients who would face the clinical decision. The clinical decision should be vital and consequential, and one in which current evidence is insufficient to inform the trade-offs of benefits versus harms associated with the different options. Researchers should make explicit what the comparators are and why they were selected, focusing on clearly describing how the chosen comparator(s) define the causal question, reduce the potential for biases, and allow direct comparisons. Generally, non-use (or no specific treatment) comparator groups should be avoided unless no specific treatment is a likely option in standard care.

Within the new phrasing of this Standard, AcademyHealth appreciated the language on interventions and comparators, specifically the mention of health care options, feeling it was sufficiently broad. However, we would suggest altering the first sentence of RQ-5—both to include additional stakeholders and to emphasize the real-world component—to read, "Interventions and comparators included in the study should correspond to the options available to patients, providers, and caregivers."

In addition, we are concerned that the many references to 'the clinical decision' exclusively results in this standard being far too limiting in its framing. Not only does PCORI use the term 'health decision' in RQ-3, straying from consistent terminology, but also, particularly at the health care organization and system levels, decisions are not solely "clinical" decisions, a term which implies a narrow focus on the actions of just the clinical provider. PCORI's priorities are far broader, and this standard should appropriately reflect that breadth for researchers; many PCORI investigations are not directly about clinical decisions but are about particular interventions with non-clinical or clinical staff engaging with patients or families. In these cases, it often makes sense to compare this to outcomes in the absence of such engaging interventions.

RQ-6: Measure outcomes that people representing the population of interest notice and care about

Identify and include outcomes the population of interest notices and cares about (e.g., survival, functioning, symptoms, health-related quality of life) and that inform an identified health decision. Define outcomes clearly, especially for complex conditions or outcomes that may not have established clinical criteria. Provide information that supports the selection of outcomes as meeting the criteria of "patient-centered" and "relevant to decision makers," such as patient and decision-maker input from meetings, surveys, or published studies. Select outcomes that reflect both beneficial and harmful effects, based on input from patient informants and people representative of the population of interest.

General feedback on the Standards for Formulating Research Questions

2. Standards Associated with Patient-Centeredness

What Changed:

Revisions specify more clearly how patients and stakeholders are involved in both the processes for conducting the research and the dissemination of the research findings. They also instruct investigators to justify their approach to engagement.

PC-1: Engage people representing the population of interest and other relevant stakeholders in ways that are appropriate and necessary in a given research context

Include individuals affected by the condition and, as relevant, their surrogates and/or caregivers. Other relevant stakeholders may include clinicians, purchasers, payers, industry, hospitals, health systems, policy makers, and training institutions. These stakeholders may be end users of the research, or, be involved in healthcare decision making.

Examples of processes in which patients, caregivers, clinicians, and other healthcare stakeholders can be involved include but are not limited to:
Formulating research questions;
Defining essential characteristics of study participants, comparators, and outcomes;
Identifying and selecting outcomes that the population of interest notices and cares about (e.g., survival, function, symptoms, health-related quality of life) and that inform decision making relevant to the research topic;
Monitoring study conduct and progress; and
Designing/suggesting plans for dissemination and implementation activities.
PCORI's Engagement Rubric provides further guidance.

When applicable, research proposals should describe how these stakeholders will be identified, recruited, and retained. If engagement is not necessary or appropriate in these processes, explain why.

PCORI is faced in developing this new set of Standards with the challenge of being prescriptive without being overly directive. When describing the relevant stakeholders within this Standard, PCORI should consider changing the language to say, "Other relevant stakeholders may include but are not limited to clinicians, purchasers…" There are other stakeholder categories that may fall beyond the list identified—one example being community-based organizations working with health care systems—that researchers should also consider when determining the types of individuals who will be involved within the study and their approach to engagement. As currently written, PCORI's list is more exclusive than illustrative.

Beyond this, here PCORI defines the role of consumers or stakeholders in a manner that is appropriate in terms of a particular research project, but it does not speak to the fact that stakeholders should be involved in the research enterprise more broadly. There's a greater role for researchers to play in engaging consumers and patients in governance and oversight processes, beyond simply the research project. AcademyHealth's Electronic Data Methods (EDM) Forum authored a paper in 2012 (http://repository.edm-forum.org/cgi/viewcontent.cgi?article=1001&context=edm_briefs) that examines and offers insight into these issues that may be useful to PCORI for incorporating these critical concepts into its revised Methodology Standards.

In addition, the final bullet regarding PCORI's Engagement Rubric is not proportionate in level and scope with the other bullets on processes. This statement should act as a broader note within the Standard, and thereby be removed from the bulleted list. We would also raise that in the last sentence of the Standard— "If engagement is not necessary or appropriate in these processes, explain why"—it is not clear whether the intent is to note where engagement is or is not appropriate for each of the project's components or for the project overall.

PC-2: Identify, select, recruit, and retain study participants representative of the spectrum of the population of interest and ensure that data are collected thoroughly and systematically from all study participants

Research proposals and subsequent study reports should describe:

- the plan to ensure representativeness of participants;
- how participants are identified, selected, recruited, enrolled, and retained in the study to reduce or address the potential impact of selection bias;
- efforts employed to maximize adherence to agreed-on enrollment practices; and
- methods used to ensure unbiased and systematic data collection from all participants.

If the population of interest includes people who are more difficult to identify, recruit, and/or retain than other study populations (for example, individuals historically underrepresented in healthcare research such as those with multiple disease conditions, low literacy, low socioeconomic status, or poor healthcare access, as well as racial and ethnic minority groups and people living in rural areas), then specify plans to address population-specific issues for participant identification, recruitment, and retention.

PC-3: Use patient-reported outcomes when patients or people at risk of a condition are the best source of information

When patients or people at risk of a condition are the best source of information regarding outcomes of interest, then the study should employ patient-reported outcome (PRO) measures in lieu of, or in addition to, measures derived from other sources. Proposals should describe: 1) the concept(s) underlying each PRO measure (e.g., symptom or impairment) and how it is meaningful to, and noticed by, patients in the population of interest; 2) how the concept relates to the health decisions the study is designed to inform; 3) how the PRO measure was developed, including how patients were involved in the development; and 4) evidence of measurement properties including content validity, construct validity, reliability, responsiveness to change over time, and score interpretability, including meaningfulness of score changes in the population of interest with consideration of important subgroups. If these measurement properties are not known, a plan for establishing the properties must be provided. Caregiver reports may be appropriate if the patient cannot self-report the outcomes of interest. If PROs are not planned for use in the study, justification must be provided.

PC-4: Support dissemination and implementation of study results

For study results that are appropriate for dissemination and implementation, involve patients and relevant stakeholders: a) in planning for dissemination from the start of the research study; b) in creating a dissemination plan for the study indicating clinical implications; c) in working with patients or organizations to report results in a manner understandable to and usable by each target audience; and d) in identifying successful strategies for adoption and distribution of study findings to targeted patient and clinical audiences.

> *AcademyHealth was perplexed by the addition of the language qualifying study results as "appropriate for dissemination and implementation" and why such a modifier would be necessary. It is unclear when—or why—PCORI would not want to share a study's findings with appropriate stakeholder audiences.*
>
> *Finally, similar to concerns raised in previous Standards, the framing of newly added standard clause 'd' is too restrictive to be of value to improving health and health care. Specifically, study findings should be adopted and distributed beyond merely 'patient and clinical audiences.' We would recommend ending the sentence at 'findings' or broadening the listed audiences.*

General feedback on the Standards Associated with Patient-Centeredness

3. Standards for Data Integrity and Rigorous Analysis

What changed:

Revisions change the order of the standards in this category and clarify the wording. Other revisions provide guidance on documenting covariates and include guidance on registering studies with the appropriate registries. A new standard, IR-6, was added to include guidance on masking.

IR-1: A priori, specify plans for quantitative data analysis that correspond to major aims

Researchers should describe the analytic approaches that will be used to address the major research aims before analysis is undertaken. These include definitions of key exposures, endpoints, and covariates. Also identify patient subgroups of interest, plans (if any) for how new subgroups of interest will be identified or how analysis plans may be adapted based on changing needs and scientific advances, and plans for how missing data will be handled.

> *AcademyHealth appreciates Standard IR-1 and its attempt to prompt researchers to think through analytic approach issues as well as to define these issues prior to conducting a base analysis. However, we feel that this Standard is missing specified plans for robustness tests. Although PCORI mentions these later in the Standards, specifically when referring to missing data methods, robustness tests are not limited to dealing only with missing data issues. To rectify this omission, PCORI could add a simple sentence, such as, "Researchers should specify their plans for robustness and sensitivity tests in advance of doing these analyses."*

IR-2: Assess data source adequacy

In selecting data sources, researchers should ensure robust capture of exposure, outcome, and relevant covariates. When statistically adjusting for covariates or confounding factors, measurement properties of the important covariates must be considered to allow adequate adjustment.

> *PCORI's Standards should reflect a wide range of data and methodologies, and the language used in IR-2 very much implies that all data are quantitative, when qualitative data are equally and often of even more importance. Investigators should specify how their analyses will incorporate and allow for the inclusion of qualitative data or mixed methods, incorporating both qualitative and quantitative data.*

IR-3: Describe data linkage plans, if applicable

For studies involving linkage of patient data from two or more sources (including registries, data networks, and others), describe
each data source and its appropriateness, value, and limitations for addressing specific research aims;
any additional requirements that may influence successful linkage, such as information needed to match patients, selection of data elements, and definitions used; and
the procedures and algorithm(s) employed in matching patients, including the success, limitations, and any validation of the matching algorithm.

When referring to data linkage plans, it's truly the combination of data sources that matters. To this point, in bullet one, where PCORI says "each data source," AcademyHealth would recommend instead altering the language to reflect the utility and fitness of the linked dataset as a whole, such as focusing on "the appropriateness and limitations of the data linkage plan," or language to this extent.

IR-4: Document validated scales and tests

Studies should include documentation of the names of the scales and tests selected, reference(s), characteristics of the scale, and psychometric properties.

IR-5: Provide sufficient information in reports to allow for assessments of the study's internal and external validity

Reporting guidelines for specific designs can be found at the relevant stakeholders EQUATOR Network website. This website has brought together all reporting guidelines that have been developed using formal approaches, many of which have been adopted by journals, such as CONSORT (for randomized clinical trials), STARD (for diagnostic tests), STROBE (for observational studies), and SRQR and/or COREQ (studies using qualitative research). Researchers should register their studies with the appropriate registry (e.g., clinicaltrials.gov for clinical studies or observational outcomes studies) and provide complete and accurate responses to the information requested (e.g., enter the required and optional data elements for clinicaltrials.gov).

There appears to be a discrepancy between Standard IR-5 and other PCORI guidance on reporting guidelines. For example, in much of PCORI's dissemination and communication work, including in its 2015 document, "PCORI's Process for Peer Review of Primary Research and Public Release of Research Findings," PCORI includes mention of both the Registry of Patient Registries (RoPR) as a repository in which "[p]atient registries must be registered" and Health Services Research Projects in Progress (HSRProj) as the database which researchers should use to register "[m]ethodological projects and others that are not appropriate for ClinicalTrials.gov or RoPR." Yet, RoPR and HSRProj are absent from the Methodology Standards entirely. For consistency with earlier guidance PCORI should account for these sites within the Methodology Standards as well. Furthermore, given that HSRProj houses the largest collection of patient-centered outcomes research projects, AcademyHealth recommends that HSRProj be included as a primary source for observational outcomes studies.

Moreover, for easy reference and convenience for users, PCORI may wish to consider providing the URLs to the websites mentioned in this Standard.

New IR-6: Masking should be used when feasible

Masking of evaluation staff should be implemented, especially in situations for which study participant and investigator masking are not feasible. When masking is not feasible, the impact of lack of masking on the results should be discussed.

First, as written, IR-6 assumes that 'evaluation staff' do not write up the results, but this is often the case and should be addressed. Additionally, for clarification purposes within this Standard, AcademyHealth recommends that PCORI change "evaluation staff" to "data collection staff."

General feedback on the Standards for Data Integrity and Rigorous Analysis

> *AcademyHealth would like to reiterate from our past comments that research projects funded by PCORI should reflect a wide range of data and methodologies – both traditional and innovative – that support robust, practical, and timely evidence generation. This set of standards could be improved upon by including a preamble stating that PCORI is referring to all kinds of data from all kinds of methods, especially including those of qualitative and mixed methods research, but is also utilizing new thinking in causal methods, including step-wedge and factorial designs and interrupted time series and regression discontinuity statistical approaches.*

4. Standards for Preventing and Handling Missing Data

What changed:

Revisions clarify the standards and identify other forms of missingness besides patient drop out. Two standards were combined, MD-2 and MD-3, due to overlap of concepts.

MD-1: Describe methods to prevent and monitor missing data

Investigators should explicitly state potential reasons that study data may be missing. Missing data can occur from patient dropout, nonresponse, data collection problems, incomplete data sources, and/or administrative issues. As relevant, the protocol should include the anticipated amount of and reasons for missing data, plans to prevent missing data, and plans to follow up with study participants. The study protocol should contain a section that addresses steps taken in study design and conduct to monitor and limit the impact of missing data. This standard applies to all study designs for any type of research question.

MD-2: Use valid statistical methods to deal with missing data that properly account for statistical uncertainty due to missingness

Valid statistical methods for handling missing data should be pre-specified in study protocols. The reasons for missing data should be considered in the analysis. A discussion of the potential ramifications of the statistical approach to missing data on the results should be provided. The plausibility of the assumptions associated with the approach should be assessed. Statistical inference of intervention effects or measures of association should account for statistical uncertainty attributable to missing data. This means that methods used for imputing missing data should produce valid confidence intervals and should permit unbiased inferences based on statistical hypothesis tests. Bayesian methods, multiple imputation, and various likelihood-based methods are valid statistical methods to deal with missing data. Single imputation methods like last observation carried forward, baseline observation carried forward, and mean value imputation are discouraged as the primary approach for handling missing data in the analysis. If investigators do use single-based imputation methods, they must provide a compelling scientific rationale as to why the method is appropriate. This standard applies to all study designs for any type of research question.

Along with the addition of "mean value imputation" as one of the examples of handling missing data, AcademyHealth also recommends including 'hot deck imputation,' in which each missing value is replaced with an observed response from a "similar" unit.[1] For additional clarification within this standard, PCORI should consider distinguishing between imputation of outcomes versus control variables.

Finally, and more generally, we encourage PCORI to push the research community to understand and report the underlying processes of data generation. The application of missing data techniques draws from a solid understanding of these underlying processes, so that the methods employed align with the mechanisms by which the data are missing.

MD-3: Record and report all reasons for dropout and missing data, and account for all patients in reports

Whenever a participant drops out of a research study, the investigator should document the following:

- the specific reason for dropout, in as much detail as possible;
- who decided that the participant would drop out; and
- whether the dropout involves some or all types of participation.

Investigators should attempt to continue to collect information on key outcomes for participants unless consent is withdrawn. All participants included in the study should be accounted for in the report, whether or not they are included in the analyses. Describe and justify any planned reasons for excluding participants from analysis. In addition, missing data due to other mechanisms (such as nonresponse and data entry/collection) should also be well documented and handled appropriately in the analyses.

AcademyHealth encourages PCORI to make its Methodology Standards as clear as possible, so they are of greatest benefit to the researcher. PCORI's request that missing data due to other mechanisms be "well documented and handled appropriately" is exceedingly vague and may not elicit the response PCORI seeks. What does PCORI consider well documented? Handled appropriately? This Standard would be enhanced with additional clarification surrounding these points, or removal of qualifiers to simply require documentation of the reasons for missingness.

As a final point, we would reiterate the point made on Standard MD-2, that pushing the research community to understand and report the underlying processes of data generation is more important than focusing on making definitive statements about the processes used in the analysis, which depend on this external content.

MD-4: Examine sensitivity of inferences to missing data methods and assumptions, and incorporate into interpretation

Examining sensitivity to the assumptions about the missing data mechanism (i.e., sensitivity analysis) should be a mandatory component of the study protocol, analysis, and reporting. This standard applies to all study designs for any type of research question. Statistical summaries should be used to describe missing data in studies, including a comparison of baseline characteristics of units (e.g., patients, questions, or clinics) with and without missing data. These quantitative results should be incorporated into the interpretation of the study and reflected in the discussion section and possibly the abstract.

General feedback on the Standards for Preventing and Handling Missing Data

5. Standards for Heterogeneity of Treatment Effects

What changed:

Existing standards were combined to streamline overlapping content in this category. Additional revisions clarify the language.

HT-1: State the goals of HTE analyses, including hypotheses and the supporting evidence base

State the inferential goal of each HTE analysis, and explain how it is related to the topic of the research. Specify whether the HTE analysis is hypothesis-driven (sometimes denoted as confirmatory), or hypothesis-generating (sometimes denoted as exploratory). Hypothesis-driven HTE analyses should be pre-specified, based on prior evidence (described clearly in the study protocol and published paper), and supported by a clear statement of the hypotheses the study will evaluate, including how subgroups will be defined (e.g., by multivariate score or stratification), outcome measures, and the direction of the expected treatment effects.

HT-2: For all HTE analyses, provide an analysis plan, including the use of appropriate statistical methods

The study protocol should unambiguously pre-specify planned HTE analyses. Appropriate methods include, but are not limited to, interaction tests, differences in treatment effect estimates with standard errors, or a variety of approaches to adjusting the estimated subgroup effect, such as Bayesian shrinkage estimates. Appropriate methods should be used to account for the consequences of multiple comparisons; these methods include, but are not limited to, p-value adjustment, false discovery rates, Bayesian shrinkage estimates, adjusted confidence intervals, or validation methods (internal or external). A common error in HTE analyses is to claim differences in treatment effect when one subgroup shows a statistically significant treatment effect and another does not.

> *AcademyHealth first recommends addressing the topic of type II errors, which is absent from the Methodology Standards altogether. To this point, researchers who focus on or care about type II errors, might not consider p-value adjustment a credible approach.*

HT-3: Report all pre-specified HTE analyses and, at minimum, the number of post-hoc HTE analyses, including all subgroups and outcomes analyzed

Protocols and study reports must report the exact procedures used to explore HTE, including data mining or any automatic regression approaches. HTE analyses should clearly report the procedures by which subgroups were defined, and the effective number of subgroups and outcomes examined. Within each subgroup level, studies should present the treatment effect estimates and measures of variability. Pre-specified HTE analyses (hypothesis-driven) should be clearly distinguished from post-hoc HTE analyses (hypothesis-generating). Statistical power for all analyses should be reported.

General feedback on the Standards for Heterogeneity of Treatment Effects

6. Standards for Data Registries

What changed:

Revisions consolidate and clarify the language. Revisions also add further guidance on stakeholder engagement, linking to data from electronic health records and claims files, and reporting and documentation of registry materials.

DR-1: Requirements for the design of registries

Registries established for conducting PCOR must have the following characteristics.

Registry Purpose and Protocol
The purpose of the registry should be clearly defined to guide the design of key registry features including, but not limited to: the target population, the research question/s to be addressed, the data source utilized, the data elements collected, data sharing policies, and the stakeholders involved in the development and use of the registry. Participants and other key stakeholders should be engaged in registry and protocol development. Registries should aim to be user-oriented in design and function.

Data Safety and Security
Registry custodians should comply with IRB requirements, local and national laws, and where applicable, HIPAA. Registries should provide information describing type of data collection (primary or secondary source data), data use agreements, informed consent documents, data security protections, plans for maintaining data protection if the registry ends, and approaches to protecting security including risk and/or process for re-identification of participants, especially for medical or claims records.

Data Elements and Quality
Standardized data element definitions and/or data dictionaries should be used whenever possible. When creating a new registry, published literature should be reviewed to identify existing, widely used definitions of outcomes, exposure, and confounders before drafting new definitions.

When collecting primary data, conduct multi-stakeholder engagement with potential participants and data users to prioritize data collection needs. When participants support their face validity, utilize validated instruments or patient-reported outcome measures when available. If secondary data sources (e.g., electronic medical records, claims data) are utilized, describe the original purpose of the secondary data and verify the accuracy and completeness of the data, as well as the approach to and validity of the linkages performed between the primary and secondary sources.

The specifics of the quality assurance plan will depend on the type of data (primary or secondary) collected by the registry. In general, the plan should address: 1) structured training tools for data abstractors/curators; 2) use of data quality checks for ranges and logical consistency for key exposure and outcome variables and covariates; and 3) data review and verification procedures, including source data verification plans (where feasible and appropriate), and validation statistics focused on data quality for the key exposure and outcome variables and key covariates. A risk-based approach to quality assurance is advisable, focused on variables of greatest importance.

Confounding
Registries should identify important potential confounders pertinent to the purpose and scope of the research during the planning phase and collect reasonably sufficient data on these potential confounders to facilitate the use of appropriate statistical techniques during the analysis phase. When conducting analysis, refer to the PCORI Methodology Standards for Data Integrity and Rigorous Analyses and the Standards for Causal Inference Methods.

Systematic Participant Recruitment and Enrollment
Develop a sampling plan (population-based or otherwise) of the target population and identify recruitment strategies for participants that minimize the impact of selection bias. Participants should be enrolled systematically, with similar procedures implemented at all participating sites and for each intervention of interest. Confirm adherence to agreed-upon enrollment practices.

Participant Follow-Up
The objective(s) of the registry should determine the type, extent, and length of participant follow-up.

Describe the frequency with which follow-up measures will be ascertained, consider linkage with other data sources such as the National Death Index to enhance long-term follow-up, and identify the date of last contact with the participant in existing registries, where appropriate. Ensure that the participants are followed in as unbiased a manner as possible, using similar procedures at all participating sites.

Monitor loss to follow-up to ensure best efforts are used to achieve follow-up time that is adequate to address the main objective. At the outset of the registry, develop a retention plan that documents when a participant will be considered lost to follow-up and what actions will be taken to minimize loss of pertinent data. Retention efforts should be developed with stakeholders to ensure the efforts are suitable for the target population and anticipated challenges are addressed appropriately.

> *First, in order to make a distinction between engagement in designing the registry infrastructure and specific studies, AcademyHealth would recommend modifying the second sentence of the "Registry Purpose and Protocol" language slightly to read, "Participants and other key stakeholders should be engaged in registry design and study protocol development."*
>
> *Furthermore, we recommend adjusting the language in first sentence of the "Data Safety and Security" section as follows: "Registry custodians should comply with IRB requirements, the HIPPA Privacy Rule, and all other applicable state and federal laws."*
>
> *Finally, unless further classification is given to the "Systematic Participant Recruitment and Enrollment" section on which sampling plans could be categorized as "otherwise," we would recommend simply dropping the phrase "(population-based or otherwise)" from the first sentence.*

DR-2: Documentation and reporting requirements of registry materials, characteristics and bias

Clearly describe, document with full citations where appropriate, and make publicly available registry materials including but not limited to registry protocols; data sharing policies; operational definitions of data elements; survey instruments utilized and patient-reported outcomes captured. Modifications to any documents or data collection instruments should be clearly described and made available for registry users and participants. Characteristics of the participants in the registry should be described. Identify how the participants may differ from the target population to help assess potential selection biases. Document loss to follow-up and describe impact on the results, utilizing sensitivity analyses (pre-specified where possible) to quantify possible biases. Report the extent of bias clearly to stakeholders who may want to utilize the registry resource.

DR-3: Adapting Established Registries for PCORI Research

Previously established registries that intend to support new clinical research may not always have been guided by all applicable Methodology Standards. When new research will use such registries, investigators should engage key stakeholders, including registry participants, to ensure:

- Informed consent documents are appropriately tailored to participant needs, characteristics and conditions
- Data elements are meaningful and useful to researchers and participants
- Recruitment and retention strategies are feasible and effective
- Registry policies are patient-centered and the use of registry data is transparent to participants
- Dissemination practices of PCORI-funded research using registry data are appropriate and effective at reaching the communities from which the data are collected
- Opportunities for bi-directional benefit between participants and researchers
- Registry materials, described in DR-2, and informed consent forms are publicly available in accessible formats

DR-4: Documentation requirements when using registry data

Researchers planning PCOR studies relying on registries must ensure that these meet the requirements contained in Standards DR-1 and DR-2 and must document each required feature of the registry(s) to be used (e.g., in an appendix to the funding application or study protocol). Deviations from the requirements with Standards DR-1 and DR-2 should be well documented and limitations of research related to the deviations from requirements should be addressed along with study findings.

General feedback on the Standards for Data Registries

> *AcademyHealth recommends that PCORI encourage both existing registries and those in development to submit a registry profile to AHRQ's Registry of Patient Registries (RoPR) to promote collaboration and encourage transparency among registry developers and users. This also represents another inconsistency with PCORI's guidance to researchers on registering studies; PCORI's dissemination and communication guidance suggests that investigators working with registries should report to RoPR, but this registry isn't referenced within the Methodology Standards.*

7. Standards for Data Networks as Research-Facilitating Structures

What changed:

Revisions provide more complete language on data integration and ownership of data.

DN-1: Requirements for the design and features of data networks

Data networks established for conducting PCOR must have the following characteristics to facilitate valid, useable data and to ensure appropriate privacy, confidentiality, and intellectual property protections:

- Data Integration Strategy—In order for equivalent data elements from different sources to be harmonized (treated as equivalent), processes should be created and documented that describe the quality and completeness of the data integration. Processes should also be created and documented that either 1) transform data elements prior to analysis or 2) make transformation logic available that can be executed when data are extracted. The selected approach should be based on an understanding of the research domain of interest.
- Risk Assessment Strategy—If data are exchanged between data partners, data custodians should develop policies for the management of the risk of use of the data other than the agreed-upon use. This should include agreements for how data will be handled and how time limits on the data will be enforced.
- Identity Management and Authentication of Individual Researchers—Develop reliable processes for verifying credentials of researchers who are granted access to a distributed research network and for authenticating them.|
- Intellectual Property Policies—A research network should develop policies for the handling and dissemination of intellectual property (IP); networks should also have an ongoing process for reviewing and refreshing those policies. IP can include data, research databases, papers, reports, patents, and/or products resulting from research using the network. Guidelines should balance 1) minimizing impediments to innovation in research processes, 2) determining whether or how IP belongs to the patients or research participants, and 3) making the results of research widely accessible, particularly to the people who need them the most.
- Standardized Terminology Encoding of Data Content—The data contents should be represented with standardized terminology systems to ensure that their meaning is unambiguously and consistently understood by parties using the data.
- Metadata Annotation of Data Content—Semantic and administrative aspects of data contents should be annotated with a set of metadata items. Metadata annotation helps to correctly identify the intended meaning of a data element and facilitates an automated compatibility check among data elements.
- Common Data Model—Individual data items should be assembled into a contextual environment that shows close or distant association among data. A common data model (CDM) specifies necessary data items that need to be collected and shared across participating institutes, clearly represents these associations and relationships among data elements, and promotes correct interpretation of the data content.

With respect to this Standard, AcademyHealth would recommend that PCORI make a few modest changes and address the following points in need of clarification:

Within the "Data Integration Strategy" bullet, in clause 2, PCORI's request for researchers to "make transformation logic" available is not so easily done. Often, it involves individuals needing both the code to transform the data as well as significant process documentation to define mapping strategies. Although this isn't a simple process, a note from PCORI within this Standard to specify what it means by 'logic' could be helpful for researchers undertaking this process.

Next, we would ask PCORI for additional illumination on the following bullet regarding "Risk Assessment Strategy." AcademyHealth assumes the greatest issue on this point is the handling risk of personal health information being released. However, in such a case, these issues are addressed in a data use agreement (DUA). Barring the DUA case, what is PCORI's threshold for a 'policy?' If PCORI is merely specifying that a DUA should be in place, and that it includes the aforementioned issues, we recommend that PCORI be precise in the Standard language.

Finally, when PCORI refers to the "standardized terminology systems" in the fifth bullet, it's unclear to which system(s) PCORI is referring—the ONC-endorsed standardized terminologies or some other resource providing standardized nomenclature? There is a vast array of standardized terminologies in the health care ether, and more specific guidelines from PCORI on its preferred set of terminology standards—whether SNOMED, LOINC, or RxNORM—would be useful.

DN-2: Selection and use of data networks

Researchers planning PCOR studies relying on data networks must ensure that these networks meet the requirements contained in Standard DN-1, and they must document each required feature of the data network(s) to be used (e.g., in an appendix to the funding application or study protocol). Deviations from the requirements should be justified by explaining why a required feature is not feasible or not necessary to achieve the overall goals of Standard DN-1.

General feedback on the Standards for Data Networks as Research-Facilitating Structures

8. Standards for Causal Inference Methods

What changed:

Revisions clarify and streamline language of the Standards. Additionally, the revisions are inclusive of non-time and time-varying exposures, specify that propensity scores are used to control for measured confounders, and specify that instrumental variables are used to address unmeasured confounding. There is also a new Standard, CI-1, requiring researchers to specify their causal model.

CI-1: CI-Model: Specify the causal model underlying the research question

Researchers should describe the causal model relevant to the research question, which should be informed by the PICOTS framework: populations, interventions, comparators, outcomes, timing, and settings. The causal model represents the key variables; the known or hypothesized relationships among them, including the potential mechanisms of effect; and the conditions under which the hypotheses are to be tested. Researchers should use the causal model to determine whether and how the study can handle bias and confounding and the extent to which valid estimates of the effects of an intervention can be generated based on the particular data source, hypothesis, and study design.

CI-2: Define and appropriately characterize analysis population used to generate effect estimates

Decisions about whether patients are included in an analysis should be based on information available at each patient's time of study entry in prospective studies or on information from a defined time period prior to the exposure in retrospective studies. For time-varying treatment or exposure regimes, specific time points should be clearly specified and relevant variables measured at baseline and up to, but not beyond, those time points should be used as population descriptors. When conducting analyses that in some way exclude patients from the original study population, researchers should describe the final analysis population that gave rise to the effect estimate(s).

> *Since the Methodology Standards revolve around patient-centered studies, the first sentence of CI-2 reads as peculiar. AcademyHealth assumes PCORI is referring to decisions about including 'specific patients,' concerning their inclusion or exclusion in studies, but PCORI may wish to stipulate its meaning more plainly.*

CI-3: Define with the appropriate precision the timing of the outcome assessment relative to the initiation and duration of exposure

To ensure that an estimate of an exposure or intervention effect corresponds to the question that researchers seek to answer, the researchers must precisely define, to the extent possible, the timing of the outcome assessment relative to the initiation and duration of the exposure.

CI-4: Measure potential confounders before start of exposure and report data on potential confounders with study results

In general, variables for use in confounding adjustment (either in the design or analysis) should be ascertained and measured prior to the first exposure to the interventions (or intervention) under study. If confounders are time varying, specific time points for the analysis of the exposure effect should be clearly specified and the confounder history up to, and not beyond, those time points should be used in that analysis.

CI-5: Report the assumptions underlying the construction of propensity scores and the comparability of the resulting groups in terms of the balance of covariates and overlap

When conducting analyses that use propensity scores to adjust for measured confounding, researchers should assess the overlap and balance achieved across compared groups with respect to potential confounding variables.

CI-6: Assess the validity of the instrumental variable (i.e., how the assumptions are met) and report the balance of covariates in the groups created by the instrumental variable

When an instrumental variable (IV) approach is used to address potential unmeasured confounding, empirical evidence should be presented describing how the variable chosen as an IV satisfies the three key properties of a valid instrument:

- the IV influences choice of the intervention or is associated with a particular intervention because both have a common cause;
- the IV is unrelated to patient characteristics that are associated with the outcome; and
- the IV is not otherwise related to the outcome under study (i.e., it does not have a direct effect on the outcome apart from its effect through exposure).

General feedback on the Standards for Causal Inference Methods

> *AcademyHealth applauds PCORI's efforts to cover a wide range of methodologies. However, after reviewing the Standards—and Standard 8 in particular—one is left with the false impression that methods employing instrumental variables and propensity scores are the primary observational data methods. AcademyHealth recommends that PCORI describe other methods such as difference-in-differences, regression discontinuity, factorial experiments and partial factorial experiments, interrupted time series, and sample selection models to give the reader a flavor for the variety of methods that are now available and are likely to be expanded in the future.*

9. Standards for Adaptive Trial Designs

What changed:

Revisions streamline and clarify the language of the Standards. Additionally, the revisions keep the reference to the 2010 CONSORT statement up to date.

AT-1: Specify planned adaptations, decisional thresholds, and statistical properties of those adaptations

The adaptive clinical trial design must be prospectively planned and the design must be clearly documented in the study protocol before trial enrollment begins, including at a minimum:

- All potential adaptations, including timing;
- Interim trial findings that will be used in determining each adaptation;
- Statistical models and decisional thresholds to be used; and
- Planned analyses of the trial endpoint(s).
- The description of the design should be sufficiently detailed that it could be implemented from the description of procedures. This specification should include a statistical analysis plan (SAP) in which all necessary detail is provided regarding planned interim and final analyses. Additionally, the statistical properties of adaptive clinical trial designs should be thoroughly investigated over the relevant range of important parameters or clinical scenarios (e.g., treatment effects, accrual rates, delays in the availability of outcome data, dropout rates, missing data, drift in participant characteristics over time, subgroup-treatment interactions, or violations of distributional assumptions). Statistical properties to be evaluated should include Type I error, power, and sample size distributions, as well as the precision and bias in the estimation of treatment effects.

AT-2: Specify the structure and analysis plan for Bayesian adaptive randomized clinical trial designs

If a Bayesian adaptive design is proposed, the Bayesian structure and analysis plan for the trial must be clearly and completely specified. This should include any statistical models used either during the conduct of the trial or for the final analysis, prior probability distributions and their basis, utility functions associated with the trial's goals, and assumptions regarding exchangeability (of participants, of trials, and of other levels). Specific details should be provided as to how the prior distribution was determined and if an informative or non-informative prior was chosen. When an informative prior is used, the source of the information should be described. If the prior used during the design phase is different from the one used in the final analysis, then the rationale for this approach should be indicated. Computational issues, such as the choice of software, the creation and testing of custom software, and software validation, should be addressed as well. Software used for Bayesian calculations during trial design, trial execution, and final analysis must be functionally equivalent. When feasible, software or programs should be made available to relevant stakeholders for evaluation and validation.

AcademyHealth commends PCORI's Standard AT-2 on the Bayesian trial structure, namely the request that researchers provide specific details about how the prior distribution was determined and if an informative or non-informative prior was chosen. To further improve this Standard, however, AcademyHealth would recommend a few minor changes:

First, in the sentence regarding computational issues, we recommend changing "be addressed as well" to "specified" to make PCORI's request more explicit. Furthermore, since the items that follow this sentence (e.g., software used for Bayesian calculations during trial design and trial execution) are very specific, appropriate requirements, PCORI should consider enumerating them clearly as a bulleted list of documentation requirements. This would also provide an additional sense of consistency with other reporting requirements included in the Methodology Standards.

AT-3: Ensure that clinical trial infrastructure is adequate to support planned adaptation(s) and independent interim analyses

The infrastructure and processes for trial implementation must be able to support the planned adaptation. The study plan must clarify who will perform the analyses to inform adaptation while the study is ongoing and who will have access to the results. The interim analyses should be performed and reviewed independent from the investigators who are conducting the trial. Trial investigators should remain blinded to changes in treatment allocation rates as this information provides data regarding treatment success. Trials with more complicated requirements, such as frequent interim analyses, require thorough testing prior to trial initiation. Such testing should involve the trial's data collection and data management procedures, the implementation of the adaptive algorithm, and methods for implementing the resulting adaptation(s). The impact on the trial's operating characteristics of delays in collecting and analyzing available outcome data should be assessed.

AT-4: When reporting adaptive randomized clinical trials, use the CONSORT statement, with modifications

The following sections of the 2010 CONSORT statement can be used to report key dimensions of adaptation:

- Adaptation of randomization probabilities (sections 8b and 13a);
- Dropping or adding study arms (sections 7b and 13a);
- Interim stopping for futility and superiority (sections 7b and 14b);
- Sample size re-estimation (sections 7a and 7b);
- Transitioning of stages (e.g., seamless Phase II/III designs) (sections 3a, 7a, 7b, and 16); and
- Modification of inclusion and exclusion criteria (sections 4a and 13a).
- CONSORT sections 16, 20, and 21 may also be expanded to report additional aspects of an adaptive trial. If the trial incorporates adaptations other than those listed above, the authors should use their judgment as to where in the CONSORT structure to include both design details and the associated results. All possible adaptations included in the prospective design, even if they did not occur, should be included in the report.

In bullet three, AcademyHealth recommends that PCORI add "or adverse outcomes" after the full "Interim stopping" clause.

General feedback on the Standards for Adaptive Trial Designs

10. Standards for Studies of Diagnostic Tests

What changed:

Revisions streamline and clarify the language, sometimes by combining existing standards. The content of the standards remains essentially the same.

DT-1: Specify clinical context and key elements of the diagnostic test

A comparative evaluation of diagnostic tests should specify each of the following items and provide rationale in support of the particular choices:

- the intended use of the test and the corresponding clinical context, including referral for additional testing, referral for additional treatments, and modification of current treatment and target populations;
- the goal of the comparison;
- the technical specifications of the tests as implemented in the study;
- the approach to test interpretation;
- the sources and process for obtaining reference standard information, when applicable;
- the procedures for obtaining follow-up information and determining patient outcomes, when applicable; and
- the clinical pathways involving the tests and the anticipated implications of test use on downstream processes of care and patient outcomes.

These items ought to be specified for all designs, including observational designs (e.g., those using medical records or registries). If these items are not available directly, validated approaches to approximating these study elements from available data should be used.

DT-2: Assess the effect of factors known to affect diagnostic performance and outcomes

Studies of diagnostic tests should include an assessment of the effect of important factors known to affect test performance and outcomes, including the threshold for declaring a "positive" test result, the technical characteristics of the test and the interpreter, and the setting of care.

DT-3: Focus studies of diagnostic tests on patient-centered outcomes, using rigorous study designs with preference for randomized controlled trials

Studies related to diagnostic and therapeutic outcomes of diagnostic testing should use a prospective randomized study design when possible. If a non-randomized design is proposed, a rationale for using an observational study (or modeling and simulation) should be provided, and efforts to minimize confounding documented.

General feedback on the Standards for Studies of Diagnostic Tests

11. Standards for Systematic Reviews

What changed:

There are no revisions to this category of standards.

SR-1: Adopt the Institute of Medicine (IOM) standards for systematic reviews of comparative effectiveness research, with some qualifications

Systematic reviews are used to answer questions based on comprehensive consideration of all the pertinent evidence, and can also identify the gaps in evidence and how they might be resolved. Standards for systematic reviews are currently in use, but credible authorities, such as Cochrane and the Agency for Healthcare Research and Quality (AHRQ), vary somewhat in their recommended standards. The IOM recently issued standards that draw broadly from available sources. The PCORI Methodology Committee endorses these standards, but recognizes that there can be flexibility in the application of some standards without compromising the validity of the review, specifically:

- Searches for studies reported in languages other than English are not routinely recommended, but may be appropriate to some topics;
- Dual screening and data abstraction are desirable, but fact-checking may be sufficient. Quality control procedures are more important than dual review per se; and
- Independent librarian peer review of the search strategy is not required; internal review by experienced researchers is sufficient.
- IOM (Institute of Medicine). Finding What Works in Health Care: Standards for Systematic Reviews. Washington, DC: The National Academies Press, 2011.

General feedback on the Standards for Systematic Reviews

12. Standards on Research Designs Using Clusters

What changed:

There are five completely new standards in this new category.

RC-1: Specify whether the study objectives, the interventions, and the primary outcomes pertain to the cluster level or individual level

Describe the target population of clusters and individuals to which the study findings will be generalizable.

Describe the clusters to be randomized and the subjects to be enrolled in the trial.

> *As a whole Standard 12, and in particular RC-1 and RC-2, is restrictive in its binary classification of the cluster and individual level of study objectives, interventions, and primary outcomes. This hierarchical manner of thinking is restrictive for research designs.*

RC-2: Justify the choice of cluster randomization

Describe the benefits and disadvantages of cluster randomization versus individual-level randomization for the proposed research. Cluster randomization should be substantiated by a sound theoretical and conceptual framework that describes the hypothesized causal pathway. Cluster randomization generally is applicable when*:

- An intervention is delivered at the cluster level
- An intervention changes the physical or social environment
- An intervention involves group processes, or
- An intervention cannot be delivered without a serious risk of contamination

*Logistical considerations can also justify cluster randomization, for example, to reduce costs or to improve participation, adherence, or administrative feasibility.

RC-3: Power and sample size estimates must use appropriate methods to account for the dependence of observations within clusters, and the degrees of freedom available at the cluster level

The methods used to reflect dependence should be clearly described. Sources should be provided for the methods and for the data used to estimate the degree of dependence. Sensitivity analyses incorporating different degrees of dependence must be reported.
For simpler designs, the dependence in the data can be reflected in the intraclass correlation.
Dependence can also be reflected in variance components
Other factors that affect the power calculation include: the design of the study, the magnitude of the hypothesized intervention effect, the pre-specified primary analysis, and the desired Type I error rate.

RC-4: Data analyses must account for the dependence of observations within clusters regardless of its magnitude

Data analyses must also reflect the degrees of freedom available at the cluster level. Investigators must propose appropriate methods for data analyses with citations and sufficient detail to reproduce the analyses.

RC-5: Because cluster randomization trials often involve a limited number of groups or clusters, stratified randomization is recommended

Non-randomized intervention trials often involve a limited number of groups or clusters, and efforts should be made to balance treatment or study conditions on potential confounders.
The recommended stratification factors are those that are expected to be strongly correlated with the outcome or with the implementation of the intervention, such as:

- Baseline value of the outcome variable
- Cluster size
- Geographic area

The PCORI Methodology Standards overall and Standard 12 (in particular Standard RC-5), would be strengthened by mentioning the importance of assessing and documenting context (which may change over time) in evaluating and comparing interventions, including the internal and external contexts. Research may be improved upon by documenting and learning from heterogeneity of results rather than simply seeking to adjust away such variation.

Furthermore, measurement of implementation factors, such as fidelity, adaptation, implementation procedures, and deviations from the planned approach, is critical in order to learn what works best for whom and in what context. Attention should be paid to how the investigator will explore the potential reasons surrounding why a seemingly good intervention fails (should that be the finding) or why some programs sites are more successful than others. Researchers should describe their approach to gathering this information—both quantitative and qualitative—on implementation and how they will integrate it with their analysis of program effects. This is also an area where qualitative and mixed methods approaches are critical to understanding the implications and sustainability of program effects.

General feedback on the Standards on Research Designs Using Clusters

AcademyHealth appreciates the addition of this new Standard, which will be increasingly important as the use of cluster design increases. This Standard is unique in that it's limited to a design-specific set of standards, while the others are somewhat design agnostic. Nevertheless, while we agree these Standards are important to include, Standard 12 includes information at a comparatively granular level. Furthermore, and notably, AcademyHealth implores PCORI to remember that cluster design is just one approach being used in the growing number of comparative studies of complex interventions. We urge PCORI to include other designs for evaluating complex interventions—including designs from implementation science—in a future iteration of the Methodology Standards.

General comments on the New and Proposed Revisions to the PCORI Methodology Standards

Comments on the New and Proposed Revisions to the PCORI Methodology Standards

The new proposed Methodology Standards continue to prove a valuable contribution to the field of health services research and to researchers wrestling with how to conduct high quality and relevant patient-centered outcomes research (PCOR). AcademyHealth appreciates the Methodology Committee's updating of these Standards to guide the field at a time of many changes in the research enterprise.

PCOR—like all health services research—has great potential to improve health, but only when it focuses on relevant questions, is produced rigorously, and is disseminated and used widely, rapidly, and by patients, caregivers, and other stakeholders. It is essential that the best scientific practices be applied in order to generate trustworthy evidence. The proposed revisions to the Methodology Standards are useful, but reflect a paradigm that is unduly limited to research on discrete clinical services and interactions. Given that PCORI's established priorities include assessing systems and eliminating disparities, PCORI needs to consider broadening the paradigm under which the standards are developed. Below are some general thoughts on how AcademyHealth feels the Standards could be improved upon:

The Methodology Report would be improved by discussing in more detail the opportunities and rigor of delivery system science, also known as improvement, implementation, or health care delivery science and of embedded research considerations as well as of methodologies beyond traditional trial methodologies (including statistical process control, step-wedge and factorial designs, and new efforts to understand rapid-cycle evaluation, such as the CMS Innovation Center has been using).

Additionally, as echoed within many comments throughout the various Standards, AcademyHealth would like to reiterate that research should reflect a wide range of data and methodologies – both traditional and innovative – that support robust, practical, and timely evidence generation. As our previous recommendations state, the Methodology Standards, which are important and complex, could be improved not only by including a descriptive paragraph per category of Standards, but also by reminding researchers that the data and methods to which it is referring are both qualitative and quantitative in nature.

Endnotes

1. Andridge, Rebecca R. and Little, Roderick J.A., "A Review of Hot Deck Imputation for Survey Non-response," Int. Stat Rev. 2010 Apr; 78(1): 40-64. Doi: 10.1111/j.1751-5823.2010.00103.x