The PCORI Methodology Standards were updated in May 2017. See the updated version of the Standards here.

The PCORI Methodology Report

Appendix A: Methodology Standards

November 2013

This page intentionally left blank.

APPENDIX A: PCORI METHODOLOGY STANDARDS

Cross-Cutting Standards for PCOR

1: Standards for Formulating Research Questions

RQ-1 Identify gaps in evidence

Gap analysis and 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.

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: 1) the specific health decision the research is intended to inform; 2) the specific population for whom the health decision is pertinent; and 3) how study results will inform the health decision.

RQ-4 Identify and assess participant subgroups

In designing studies, researchers should identify participant subgroups of interest and, where feasible, design the study with adequate precision and power to reach conclusions specific to these subgroups.In addition, subgroup information should be reported for later systematic reviews.

RQ-5 Select appropriate interventions and comparators

When evaluating an intervention, the comparator treatment(s) must be chosen to enable accurate evaluation of effectiveness or safety compared to other viable options for similar patients. Researchers should make explicit what the comparators are and how 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.

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, function, 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 based on input directly elicited from patient informants and people representative of the population of interest, either in previous studies or in the proposed research.

2: Standards Associated with Patient-Centeredness

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

People representing the population of interest include individuals who have the condition or who are at risk of the condition and, as relevant, their surrogates or caregivers. Other relevant stakeholders may include clinicians, administrators, policy makers, or others involved in healthcare decision making. Stakeholders can be engaged in the processes of:

- 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.

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.

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: 1) the plan to ensure representativeness of participants; 2) how participants are identified, selected, recruited, enrolled, and retained in the study to reduce or address the potential impact of selection bias; 3) efforts employed to maximize adherence to agreed-on enrollment practices; and 4) 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-unique 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

Support dissemination and implementation of study results by suggesting strategies, indicating clinical and policy implications, and working with patients or organizations to report results in a manner understandable to each target audience.

3: Standards for Data Integrity and Rigorous Analyses

IR-1 Assess data source adequacy

In selecting variables for confounding adjustment, researchers should assess the suitability of the data source in terms of its ability to assure robust capture of needed covariates.

IR-2 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 1) each data source and its appropriateness, value, and limitations for addressing specific research aims; 2) any additional requirements that may influence successful linkage, such as information needed to match patients, selection of data elements, and definitions used; and 3) the procedures and algorithm(s) employed in matching patients, including the success, limitations, and any validation of the matching algorithm.

IR-3 A priori, specify plans for data analysis that correspond to major aims

Researchers should describe the analytic approaches that will be used to address the major research aims prior to data collection. 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.

IR-4 Document validated scales and tests

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

IR-5 Use sensitivity analyses to determine the impact of key assumptions

The results of these sensitivity analyses should be reflected in the interpretation of results.

IR-6 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 EQUATOR Network website (www.equator-network.org). 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), and STROBE (for observational studies).

4: Standards for Preventing and Handling Missing Data

MD-1 Describe methods to prevent and monitor missing data

Investigators should explicitly anticipate potential problems of missing data. The study protocol should contain a section that addresses missing data issues and steps taken in study design and conduct to monitor and limit the impact of missing data. Missingness can occur from patient dropout, failure to provide data, and/or administrative or data management issues. As relevant, the protocol should include the anticipated amount of and reasons for missing data, as well as plans to follow up with participants. This standard applies to all study designs for any type of research question.

MD-2 Describe statistical methods to handle missing data

Statistical methods for handling missing data should be pre-specified in study protocols. The reasons for missing data should be considered in the analysis. The plausibility of the assumptions associated with the approach should be assessed. A discussion of the potential ramifications of the approach to missing data on the results should be provided. This standard applies to all study designs for any type of research question.

MD-3 Use validated methods to deal with missing data that properly account for statistical uncertainty due to missingness

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 have valid Type I error rates and that confidence intervals should have the nominal coverage properties. This standard applies to all study designs for any type of research question. Bayesian methods and methods such as multiple imputation satisfy this condition, along with various likelihood-based and other validated methods. Single imputation methods like last observation carried forward and baseline observation carried forward 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.

MD-4 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: 1) the specific reason for dropout, in as much detail as possible; 2) who decided that the participant would drop out; and 3) 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. This standard applies to all prospective study designs that aim to assess intervention effectiveness. All participants included in the study should be accounted for in the report, whether or not they are included in the analysis. Describe and justify any planned reasons for excluding participants from analysis.

MD-5 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.

5: Standards for Heterogeneity of Treatment Effects

HT-1 State the goals of HTE analyses

State the inferential goal of each HTE analysis, specifying how it is related to the topic of the research, translate this into an analytic approach, and highlight the linkages between the two. Identify analyses as hypothesis driven (sometimes denoted confirmatory), or hypothesis generating (sometime denoted exploratory).

HT-2 For all HTE analyses, pre-specify the analysis plan; for hypothesis-driven HTE analyses, pre-specify hypotheses and supporting evidence base

The study protocol should unambiguously pre-specify planned HTE analyses. Pre-specification of hypothesis-driven HTE analyses should include a clear statement of the hypotheses the study will evaluate, including how groups will be defined (e.g., by multivariate score or stratification) and outcome measures, and the direction of the expected treatment effects. The pre-specified hypotheses should be based on prior evidence, which should be described clearly in the study protocol and published paper.

HT-3 All HTE claims must be based on appropriate statistical contrasts among groups being compared, such as interaction tests or estimates of differences in treatment effect

A common error in HTE analyses is to claim differences in treatment effect when one group shows a statistically significant treatment effect and another does not. To claim differences in treatment effect among subgroups, appropriate statistical methods must be used to directly contrast them. Such contrasts 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. Within each subgroup level, studies should present the treatment effect estimates and measures of variability.

HT-4 For any HTE analysis, report all pre-specified analyses and, at minimum, the number of post hoc 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 (e.g., by categorical predictors or continuous risk scores) and the effective number of subgroups and outcomes examined. If a non-prespecified stratum or subgroup is claimed to show a treatment effect that is different from others, methods should be used that account for the number of contrasts examined. These methods include, but are not limited to, p-value adjustment, false discovery rates, Bayesian shrinkage estimates, adjusted confidence intervals, and validation methods (internal or external).

Standards for Specific Study Designs and Methods

6: Standards for Data Registries

DR-1 Requirements for the design and features of registries

Registries established for conducting PCOR must have the following characteristics to facilitate the collection and aggregation of usable data, to ensure appropriate privacy and confidentiality, to document changes to the registry protocol, and to guide robust analyses that include important confounders.

A. Patient Follow-up
The objective(s) of the registry should determine the type, extent, and length of patient follow-up. Describe what triggers the follow-up, the follow-up measures, and the last contact with the patient. Ensure that the planned follow-up time is adequate to address the main objective and that planned patient-retention efforts are suitable to the target population and anticipated challenges. Describe expected loss to follow-up and potential effect on the results, including possible biases resulting from differential loss.

B. Data Safety and Security
Registry custodians should provide transparency for institutional review boards by describing data use agreements, informed consent, data security, and approaches to protecting security including risk of re-identification of patients. If using previously collected data, describe how these address the risk of re-identification of patients and the actual use of data compared with the originally designed and consented use of the data.

C. Data Quality Assurance
A quality assurance plan for registries should address: 1) structured training tools for data abstractors; 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 and validation statistics focused on the key exposure and outcome variables and covariates for which sites may be especially challenged. A risk-based approach to quality assurance is advisable, focused on variables of greatest importance.

D. Document and Explain Any Modifications to the Protocol

Modifications to a registry protocol may be necessary for a variety of reasons. When modifications are necessary, they should be explained, documented, and made available to anyone planning to use the registry data.

E. Consistent Data Collection

Clear, operational definitions of data elements should be provided. Create and distribute standard instructions to data collectors. Use standardized data element definitions and/or data dictionaries whenever possible. When creating a new registry, published literature should be reviewed to identify existing, widely used definitions before drafting new definitions.

F. Systematic Patient Enrollment and Follow-up

Enroll patients systematically and follow them in as unbiased a manner as possible, using similar procedures at all participating sites. Describe how patients and providers were recruited into the study to allow the impact of selection bias to be clearly understood; for example, by explaining whether the sampling was population-based or otherwise and any efforts employed to confirm the quality of adherence to agreed-on enrollment practices.

G. Monitor and Minimize Loss to Follow-up

Monitor loss to follow-up to ensure that follow-up is reasonably complete for the main objective. Minimizing loss to follow-up requires having a target and advance planning for what actions will be employed in the event that this target is in jeopardy. At the outset of the registry, develop a patient retention plan that documents when a patient will be considered lost to follow-up and what actions will be taken to minimize such loss. At the enrollment visit, consider collecting multiple types of contact information (e.g., telephone, mailing address, and email address) for the patient, as well as collecting contact information for an alternate contact if the patient cannot be reached directly. Verify contact information at each subsequent visit and update as needed. When a patient misses a visit, contact the patient following a standard protocol (e.g., phone call one day after missed visit, email one week after missed visit). If the patient withdraws from the registry, attempt to document the reason for withdrawal so that issues can be identified and addressed (e.g., overly burdensome patient-reported outcome measures). Efforts at minimizing loss to follow-up should be tempered by considerations and sensitivity to repeated intrusions on patients and to the health conditions and interventions under study. Consider collecting enough information to permit accurate linkage with other data sources, such as the National Death Index, for long-term follow-up.

H. Collect Data to Address Confounding

Registries should identify important potential confounders 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.

DR-2 Selection and use of registries

Researchers planning PCOR studies relying on registries must ensure that these meet the requirements contained in Standard DR-1 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 should be justified by explaining why a required feature is not feasible or not necessary to achieve the overall goals of Standard DR-1.

DR-3 Robust analysis of confounding factors

In studies that use registries to evaluate the comparative effectiveness or safety of interventions, investigators should select an approach for adjusting for known and measured confounders, such as multivariable regression analysis or propensity scores to create matched comparison groups, or an instrumental variable analysis if a valid instrument is available. It is also desirable to examine the robustness of the results through sensitivity analyses focused on testing key assumptions and evaluating the likely impact of unmeasured confounders. The rationale for using selected techniques, any assumptions made, and the strengths and limitations of the techniques should be described in reports of the study findings to allow for informed interpretation of the results.

7. Standards for Data Networks as Research-Facilitating Structures

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:

A. **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 either 1) transform and standardize 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.

B. **Risk Assessment Strategy—**Data custodians should measure the risk of re-identification of data and apply algorithms to ensure that the desired level of confidentiality is achieved to meet the need of the particular PCOR application.

C. **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.

D. **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 and 2) making the results of research widely accessible, particularly to the people who need them the most.

E. **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.

F. **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.

G. **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.

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.

8. Standards for Causal Inference Methods

CI-1 Define analysis population using covariate histories

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 the covariates history up to and not beyond those time points should be used as population descriptors.

CI-2 Describe population that gave rise to the effect estimate(s)

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).

CI-3 Precisely define 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 the timing of the outcome assessment relative to the initiation and duration of the exposure.

CI-4 Measure confounders before start of exposure and report data on 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 therapy (or therapies) 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 balance covariate distributions across intervention groups, 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 for all instrumental variable analyses

When an instrumental variable (IV) approach is used, empirical evidence should be presented describing how the variable chosen as an IV satisfies the three key properties of a valid instrument: 1) the IV influences choice of the intervention or is associated with a particular intervention because both have a common cause; 2) the IV is unrelated to patient characteristics that are associated with the outcome; and 3) 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).

9. Standards for Adaptive and Bayesian Trial Designs

AT-1 Specify planned adaptations and primary analysis

The adaptive clinical trial design should be prospectively planned and the design clearly documented, including:

- All potential adaptations, including timing;
- Trial results and populations that will be used in determining each adaptation;
- Statistical models to be used; and
- Planned analysis of the primary endpoint(s).

The description of the design should be sufficiently detailed that it could be implemented from the description of procedures. The specification of the design should be completed and documented in the trial protocol before enrollment begins. This specification should include, in all but the simplest designs, a statistical analysis plan (SAP) that is separate from the trial protocol in which all necessary detail is provided regarding planned interim and final analyses. Prior specification is a prerequisite for valid and meaningful evaluation of an adaptive design.

AT-2 Evaluate statistical properties of adaptive design

While not necessary for simple designs, the statistical properties of complex 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. Additional performance metrics may also be evaluated (e.g., the frequency with which specific adaptations occur, the likelihood of substantial covariate imbalance, the likely adequacy of final data for subgroup and safety analyses). The programming code used to create the simulations should be retained with version control. The programming code and software used should be made available to stakeholders who have a need to know, including reviewing agencies.

AT-3 Specify 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. Utility functions, if employed, should be defined, and their source should be described. 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.

AT-4 Ensure clinical trial Infrastructure is adequate to support planned adaptation(s)

The clinical trial infrastructure, including centralized randomization, data collection related to the assessment and recording of key outcomes, data transmission procedures, and processes for implementing the adaptation (e.g., centralized, web-based randomization), must be able to support the planned trial. In simple adaptive trials, qualitative verification of the capabilities of the proposed trial infrastructure may be adequate. 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. The study plan should 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 by an analytical group that is 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.

AT-5 Use the CONSORT statement, with modifications, to report adaptive randomized clinical trials

The following sections of the 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 criterion (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.

10. Standards for Studies of Diagnostic Tests

DT-1 Specify clinical context and key elements of diagnostic test study design

A comparative evaluation of diagnostic tests should specify each of the following items and provide rationale in support of the particular choices: 1) 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; 2) the goal of the comparison; 3) the technical

specifications of the tests as implemented in the study; 4) the approach to test interpretation; 5) the sources and process for obtaining reference standard information, when applicable; and 6) the procedures for obtaining follow-up information and determining patient outcomes, when applicable. 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 Study design should be informed by investigations of the clinical context of testing

Design of comparative effectiveness studies should outline clinical pathways involving the tests and the anticipated implications of test use on downstream processes of care and patient outcomes. In the written research methods and study protocol, investigators should give examples of clinical pathways to demonstrate thorough understanding of the clinical context.

DT-3 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-4 Structured reporting of diagnostic comparative effectiveness study results

Broadly accepted checklists for reporting studies and assessing study quality, such as CONSORT, STARD, and QUADAS, should be consulted and utilized. Consult the CONSORT 2010 checklist for reporting randomized controlled trials. Consult the STARD checklist for reporting diagnostic accuracy studies. Consult the QUADAS-2 (updated in 2011) for additional guidance on reporting information that would be more useful to systematic reviews of diagnostic accuracy studies.

DT-5 Focus studies of diagnostic tests on patient-centered outcomes, using rigorous study designs with preference for randomized controlled trials

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

11: Standards for Systematic Reviews

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.