Methodology Advisory Panel (MAP) Proposal Review
Investigator Guidelines

Methodology Advisory Panel (MAP) Overview

The PEDSnet Methodology Advisory Panel (MAP) is comprised of doctoral-level statisticians, epidemiologists, clinical trial experts, and social scientists with at least one member based at each of the eight institutions that comprise PEDSnet (click for individual profiles). The charge of the MAP is to assist investigators in developing high quality protocols that will result in actionable results for health care providers and patients or changes in practice. All studies using PEDSnet data are required to undergo a MAP review. The main purpose of the reviews is to provide methodologic guidance. All PEDSnet applicants are expected to seek statistical expertise to assist in the development of their application.

Guidelines - Goal

The goal of this document is to provide investigators with guidelines on what they can expect the MAP reviewers to focus on most closely when conducting PEDSnet concept proposal reviews, thereby maximizing efficiency and minimizing turnaround time of reviews and follow up discussions. [See also Velentgas, et al. 2013 and Fleurence, et al. 2015 for general guidance.]

Guidelines are organized by methodological sub-topic, as follows.

I. Appropriateness of the aims and hypotheses

Clearly define the questions of interest. Always consider what data are available and can support the aims and hypotheses.

- Can the data in PEDSnet answer the question of interest?
- What elements of the electronic health records can be used?
- What elements are absent?

II. Appropriateness of the study design

Questions will drive the choice of design, which in turn will drive the method of analysis.

For objective causal inference, design trumps analysis (Rubin 2008). No matter how elegant (and correct) the analysis may be, the design will drive the study. As Rubin notes, for observational studies, it might be helpful to design an observational study to mimic a randomized experiment. If there are available guidelines for study design, then it might be easier to follow them than to reinvent the design wheel. A common use of PEDSnet data might be prediction (or prognostic modeling) using available data. A recent guidance is TRIPOD, which has not only a checklist but an extensive set of comments in the accompanying Explanation and Elaboration (Collins, et al. 2015).

In these studies, determining thresholds and the relative costs of false positive and false negative decisions (or classifications) will be especially important.

III. Intervention approach (if applicable)

Intervention is yet another feature of design. Two intervention issues usually of particular interest are fidelity of the actual intervention to the protocol and adherence to the intervention.

IV. Validity and reliability of the measurement methods

Measurement error in exposures and outcomes can bias results toward the null and reduce statistical power. Measurement error of covariates can lead to bias toward or away from the null. Measurements might differ across institutions. Combining data across 8 institutions is not the same as having all the data from one site. Coding of conditions, reporting of values, and the places in which those values are recorded, can and often do vary across institutions, and this variation might affect some conditions more than others. An initial goal of a study might be to evaluate to what extent measures can be derived across different electronic health records, as they are implemented in different hospitals.

V. Sample size and statistical power

Sample size calculations can be difficult when the design is complex. Perhaps the most important question in estimating power is to distinguish among the requirements of different designs. Superiority studies (*is A better than B*) are relatively easy to estimate statistical power. Non-inferiority studies (*is A at least as good as B*) and equivalence studies (*are A and B equally good*) require quite different calculations. In any case, the study design will drive the type of power calculation. Reporting the software used will not only help the reviewers, it will also save words.

Calculations should focus on the particular endpoints of importance. Because estimates will be derived across 8 hospitals, they really should anticipate heterogeneity across the 8 sites. If there are planned subgroup analyses, or if effect modification or heterogeneity of treatment effect are questions of interest, then the power calculations should support those questions. Finally, less than complete adherence, or switching of treatment group, will often bias results toward the null. In superiority studies, that bias will reduce power. In non-inferiority and equivalence studies, bias towards the null also undermines the design.

VI. Appropriateness of data analysis plans and statistical tests including adjustment for confounders and assessing effect modification where appropriate

(a) Description

The cohort description should cover inclusion and exclusion criteria as well as the manner in which available data will support the specification of the cohort. Be skeptical from the start. Assume that this work cannot be done with the data available, and then report efforts to validate the data. Make recommendations as to what aspects of the "data generating engine" (the data systems across the PEDSnet participating institutions) might need improvement to fulfill longer terms goals.

(b) Modeling

Reporting the software used (and the program to be used) can be a helpful shorthand to describing the analytic approach. PEDSnet data are inherent multi-level, and this data structure will require attention to sources of bias and variance from the data structure.

(c) Confounding

Collection and analysis of observed factors that are correlated with unobserved factors can reduce the potential for unobserved confounding. On top of the usual concerns of confounding by patient-level factors, another common concern in multi-institutional designs is confounding by hospital cluster. If outcomes vary by hospital (over and

above variation that can be explained by patient-level covariates), and outcomes vary by hospital over and above variation explained by observed patient-level factors, then hospital can be a confounder.

(d) Effect modification (Heterogeneity of treatment effect (HTE))

Several issues fall under the topic of heterogeneity. Will there be adequate data to identify patients who are more likely to benefit from treatment or more likely to be at risk for the adverse treatment effects outcomes? Not only might outcomes vary by hospital or by patient characteristics (outcome heterogeneity), effects of exposures or treatments might also vary across hospitals or by patient subgroups (treatment heterogeneity). Identification of subgroups might require a separate statistical model. The association of interest between exposure and outcome might vary across hospitals over and above that explained by patient subgroups. Do all patient subgroups across institutions have the same association of exposure and outcome?

Of particular importance, heterogeneity is scale specific: effect modification on the relative scale as measured by relative risk, for example, does not equate to effect modification on the absolute risk difference scale (VanderWeele & Knol 2011). The biostatistics literature on patient-level factors and treatment benefit (HTE) in RCTs is now large, and possible methods for HTE abound (Lipkovich, et al. 2017). If a goal is in part to identify patients who have especially good or especially poor treatment effects, then the proposal should reference, explain, and defend the chosen statistical approach. The proposal should describe patients who can be identified for future trials or studies based on these results.

(e) Sensitivity analysis

All observational studies need sensitivity analyses for unobserved confounders (Ding, et al. 2016). Confounding can not only lead to biased point estimates; it can diminish or eliminate statistical significance and thus diminish statistical power. Results can also be sensitive to missing data and measurement error.

VII. Missing data

Every proposal needs to explain the reason for the missing data. What might cause the data to be absent in PEDSnet: institutional processes, physician habits or customs, or patient factors? It is often helpful to distinguish between missing Ys and missing Xs.

Missing outcomes (Ys): This is the classic problem of loss to follow up. Patients move, fail to keep appointments, or seek follow-up care outside of the participating institution. What procedure will be used to account for or impute missing outcomes?

Missing exposures (Xs): A record review is likely to be extremely important and helpful.

Consider a plan for collecting data related to the probability of having missing data or dropout, and correlated to the value of the data if they could be observed. See also a summary of recent guidance from the Institute of Medicine (Little, et al. 2012).

VIII. Other

Report sufficient details, and especially key references, where the methods are unique or especially applicable to the study design. Reference to software (specific program) helps to communicate methods succinctly.

References *(hyperlinked references indicate open access to full text)*

Velentgas P, Dreyer NA, Nourjah P, Smith SR, Torchia MM, eds. Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide. AHRQ Publication No. 12(13)-EHC099. Rockville, MD: *Agency for Healthcare Research and Quality*; January 2013.

Fleurence R, Whicher D, Dunham K, Gerson J, Newhouse R, Luce B. The Patient-centered Outcomes Research Institute's Role in Advancing Methods for Patient-centered Outcomes Research. *Med Care.* 2015; 53(1):2-8.

Rubin, D. For Objective Causal Inference, Design Trumps Analysis. *Ann Applied Stat.* 2008; 2(3):808-840. Link to abstract.

Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. ***Ann Intern Med.*** 2015 Jan 6;162(1):55-63.

Little RJ, D'Agostino R, Cohen ML, Dickersin K, Emerson SS, Farrar JT, et al. The prevention and treatment of missing data in clinical trials. *N Engl J Med.* 2012; 367(14):1355-60.

VanderWeele TJ, Knol MJ. Interpretation of subgroup analyses in randomized trials: heterogeneity versus secondary interventions. Ann Intern Med. 2011; 154(10):680-3.

Lipkovich I, Dmitrienko A, RB D'Agostino Sr. Tutorial in biostatistics: data-driven subgroup identification and analysis in clinical trials. *Stat Med.* 2017 Jan 15;36(1):136-196. Link to abstract.

Ding P, VanderWeele TJ. Sensitivity Analysis Without Assumptions. *Epidemiology.* 2016; 27(3):368-77.