Guiding Principles for Large-Scale Evidence Generation and Evaluation in a Network of Databases (LEGEND)

# Goal

The goal of LEGEND is to generate evidence on the effects of medical interventions using observational healthcare data to support clinical decision making.

# Main principles

1. **LEGEND will generate evidence at large-scale.**

* Instead of answering a single question at a time (e.g. the effect of one treatment on one outcome), LEGEND answers large sets of related questions at once (e.g. the effects of many treatments for a disease on many outcomes).
* **Aim**: Avoids publication bias, achieves comprehensiveness of results, and allows for evaluation of the overall coherence and consistency of the generated evidence

1. **Dissemination of the evidence will not depend on the estimated effects.**

* All generated evidence is disseminated at once.
* **Aim**: Avoids publication bias and enhances transparency

1. **LEGEND will generate evidence using a pre-specified analysis design.**

* All analyses will be decided prior to analysis execution.
* **Aim**: Avoids p-hacking

1. **LEGEND will generate evidence by consistently applying a systematic process across all research questions.**

* This principle precludes modification of analyses to obtain a desired answer to any specific question. This does not imply a simple one-size-fits-all process, rather that the logic for modifying an analysis for specific research questions should be explicated and applied systematically.
* **Aim**: Avoids p-hacking and allows for evaluation of the operating characteristics of this process (Principle 6).

1. **LEGEND will generate evidence using best practices.**

* LEGEND answers each question using current best practices, including advanced methods to address confounding such as propensity scores. Specifically, we will not employ suboptimal methods (in terms of bias) to achieve better computational efficiency.
* **Aim**: Minimizes bias

1. **LEGEND will include empirical evaluation through the use of control questions.**

* Control questions are questions where the answer is known. These allow measuring the operating characteristics of our systematic process, including residual bias. We subsequently account for this observed residual bias in our p-values, effect estimates, and confidence intervals (CIs) using empirical calibration.
* **Aim**: Enhances transparency on the uncertainty due to residual bias

1. **LEGEND will generate evidence using open source software that is freely available to all.**

* The analysis software is open to review, evaluation, and available for replicating analyses down to the smallest detail.
* **Aim**: Enhances transparency and allows replication

1. **LEGEND will not be used to evaluate new methods.**

* Even though the same infrastructure used in LEGEND may also be used to evaluate new causal inference methods, generating clinical evidence should not be performed at the same time as method evaluation. This is corollary of Principle 5, since a new method that still requires evaluation cannot already be best practice. Also, generating evidence with unproven methods can hamper interpretability of the clinical results. Note that LEGEND does evaluate how well the methods it uses perform in the specific context of the questions and data used in a LEGEND study (Principle 6).
* **Aim**: Avoids bias and improves interpretability.

1. **LEGEND will generate evidence across a network of multiple databases.**

* Multiple heterogeneous databases (different data capture processes, healthcare systems, and populations) will be used to generate the evidence to allow an assessment of the replicability of findings across sites.
* **Aim**: Enhances generalizability and uncovers potential between-site heterogeneity

1. **LEGEND will maintain data confidentiality, patient-level data will not be shared between sites in the network.**

* Not sharing data will ensure patient privacy, and comply with local data governance rules.
* **Aim**: Privacy