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. **Evidence will be generated at large-scale.**

* The main purpose of this principle is to achieve completeness, and to facilitate analysis of the overall distribution of effect size estimates. Rather than answering single questions (e.g. what is the effect of exposure E on outcome O), large sets of questions will be answered (e.g. what are the effects of exposures E1â€¦En on outcomes O1…Om)

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

* The main purpose of this principle is to avoid publication bias. No priority will be given to â€˜statistically significantâ€™ findings, nor will null findings or findings at odds with current knowledge be excluded from dissemination.

1. **The evidence will be generated using a pre-specified analysis design.**

* The main purpose of this principle is to avoid p-hacking (modifying the analysis to achieve a desired effect estimate). The analysis design should be specified in the form of a protocol as well as open source executable code, and should be publicly disclosed before executing the analysis.

1. **Evidence will be generated by consistently applying a systematic approach across all research questions.**

* The main purpose of this principle is to avoid p-hacking (modifying the analysis to achieve a desired effect estimate), and to enable the evaluation of the method used. If for a particular research question someone proposes the analysis should be modified to account for unique attributes of that research question, this proposal should be formulated as a general rule. This rule can then be implemented, and evaluated on its merits. If the evaluation shows adoption of the rule increases performance, the rule will become part of the systematic approach.

1. **The evidence will be generated using best-practices.**

* Whenever possible, best practices will be decided on based on empirical evaluation of the methods. Empirical evaluation should include enough negative and positive controls to ensure generalizability and accuracy (in other words, showing a method works in one or two examples is not considered â€˜empirical evaluationâ€™). If empirical evaluation is not yet performed, best practices may be defined by the opinions of the LEGEND Leadership.

1. **The evidence generation process will be empirically evaluated by including control research questions where the true effect size is known.**

* The main purpose of this principle is to make sure the operating characteristics of the methods and data are known. Since non-interventional studies are prone to systematic error, the potential magnitude of this systematic error should be quantified and accounted for (for example by computing calibrated confidence intervals and p-values).

1. **The evidence will be generated using open source software that is freely available to all.**

* This purpose aims to achieve full transparency of the analyses, and allow validation of the software by others.

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

* The main purpose of this principle is to keep the focus on evidence generation using best practices. LEGEND researchers can perform method evaluation outside of LEGEND to inform best practices used in LEGEND.

1. **Evidence will be generated in a network of heterogeneous databases.**

* The main purpose of this principle is to assess replicability of findings across databases. Additionally, it can provide larger sample size.

1. **No patient-level data will be shared between sites in the network, only aggregated data.**

* The main purposes of this principle is to ensure patient privacy, and comply with local data governance rules.