TuftsMedicine
Tufts Medical Center
NEWDIGS

Process for Identification & Selection of Clinical Outcomes & Impact Metrics

LEAPS METRICS Team

Purpose

The purpose of this framework is to support the identification of metrics and thresholds for predictive modeling and reimbursement of drug therapy regimens through a defined multi-stakeholder process. The two primary areas of focus are:

- **Clinical outcomes:** Condition/treatment-specific outcome measures that are acceptable to key stakeholders and are technically feasible for predictive modeling. These outcome measures may serve as inputs for Precision Reimbursement contract designs and target variables for predictive modeling

- **Impact metrics:** System-wide measures for tracking overall pilots that reflect the impact of the new integrated system capability (predictive models + payment models) on perceived benefits and risks for each stakeholder. Impact metrics are expected to be applicable to multiple disease areas.

Framework Application[1]

Table 1

Domain: Scope Specification	
LEAPS Process	**ICIs in NSCLC Use Case**
Align **Scope**: The setting(s) in which the outcomes sets are to be applied	US real-world care system
The health condition(s) covered by the outcomes sets	Advanced or metastatic non-small cell lung cancer (adv/met NSCLC)
The population(s) covered by the outcomes sets	Adult adv/met NSCLC patients (regardless of insurance coverage, practice setting, or US region)
The intervention(s) covered by the outcomes sets	Immune checkpoint inhibitors (ICIs)

Table 2

Domain: Stakeholders involved	
LEAPS Process	**ICIs in NSCLC Use Case**
Invite: Apply LEAPS stakeholder mapping process to identify those who will use the outcomes sets in practice, RWD analysis, or coverage decisions[2]	Patient, clinician, payer, developer, and analytics team representatives

Table 3

Domain: Consensus process*	
LEAPS Process	**ICIs in NSCLC Use Case**
Gather initial list of outcomes considering views of all stakeholders • Collection process: should include multi-stakeholder meeting(s), review of existing literature (both trials and RWE studies)	a) Clinical outcomes: initial list developed at multi-stakeholder Design Labs, and by Clinical Outcomes subteam, informed by lit review b) Impact Metrics: initial list developed by Impact Metrics subteam, informed by lit review
Filter initial list of metrics for feasibility / practicality / duplication • Describe (implicit and explicit) criteria used to create "short list"[3] • Note all measures ranked as important by stakeholders, but not included for practicality / lack of data	Analytics/modeling team; patient, clinician, payer, and developer representatives
Prioritize: A scoring process and consensus definition is used, described, and refined based on continuous learning process • Scoring and consensus process: modified Delphi with ranking of each option or cumulative voting[4] • Nominate proxy measures for important but technically infeasible outcomes	a) Clinical outcomes: initial list prioritization by the ICIs in NSCLC pilot team; voting exercise at November Design Lab b) Impact Metrics: initial list prioritization by the ICIs in NSCLC pilot team; voting exercise at November Design Lab
Establish **Thresholds** for action • Thresholds for decision-making (e.g., incremental difference needed) to be sought for each stakeholder category	Whiteboarding session at Design Lab with representative group of stakeholders

• Care is taken to avoid ambiguity of language used in the list of outcomes	

**After scope specification and stakeholder mapping, the consensus process is applied to both a) Disease-specific Clinical Outcomes and b) Systems-level Impact Metrics*

[1] Modified from Core Outcome Set-STAndards for Development (COS-STAD) criteria as described in Gargon, *et al.* J Clin Epidemiol 2019; 112:36e44. (Scope specification: COS-STAD items 1-4; Stakeholders involved: items 5-7; Consensus process: items 8-11.)

[2] For this pilot use case in NSCLC, NEWDIGS and LEAPS convened 2 multi-stakeholder Design Labs reaching out to stakeholders with interest in oncology topics.

[3] For both Clinical Outcomes and Impact Metrics, participants on the full METRICS team were asked to consider which outcome measure was 1) most important and 2) most feasible. Each person had one vote for each category; the measures with 0 or 1 only vote (<15% of participants) will not proceed to the Prioritization round, but will be listed for Design Lab participants along with the rationale for not including.

[4] At least one round of group voting is conducted as part of consensus process, with the first round of voting typically conducted with participants blinded to one another's votes. In cumulative voting, each voter is asked to distribute a fixed number of points (e.g., 10) among the candidate options in any way they please. The option with the most total points is the highest priority. See: Stanford Encyclopedia of Philosophy, 2019. Voting Methods. At: https://plato.stanford.edu/entries/voting-methods/. Alternately, ranking or scoring of each option by each voter is often conducted as part of an electronic survey, e.g., Likert 1-9 scales for importance, as described in Zimmerman, et al. JAMA Netw Open 2022: 5: e2233872.