Designing Screening Trials: from test sensitivity to late-stage cancer incidence reduction

Kehao Zhu; Yingqi Zhao; *Yingye Zheng*
zhukehao@uw.edu; yqzhao@fredhutch.org; yzheng@fredhutch.org

Fred Hutch Cancer Center

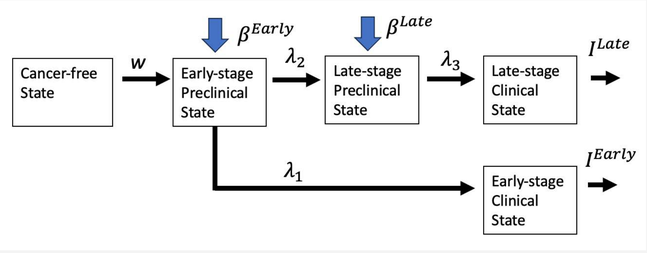
NIH NATIONAL CANCER INSTITUTE Early Detection Research Network

Aims

- Main Aim: design a Phase 4 clinical utility trial
- Primary endpoint: reduction in the late-stage cancer incidence
- Task 1: project the time-varying effect size based on the information on test characteristics
- Task 2: recommend optimal design in term of statistical power
- Secondary Aim : define test performance criteria for Phase 3 studies

Method

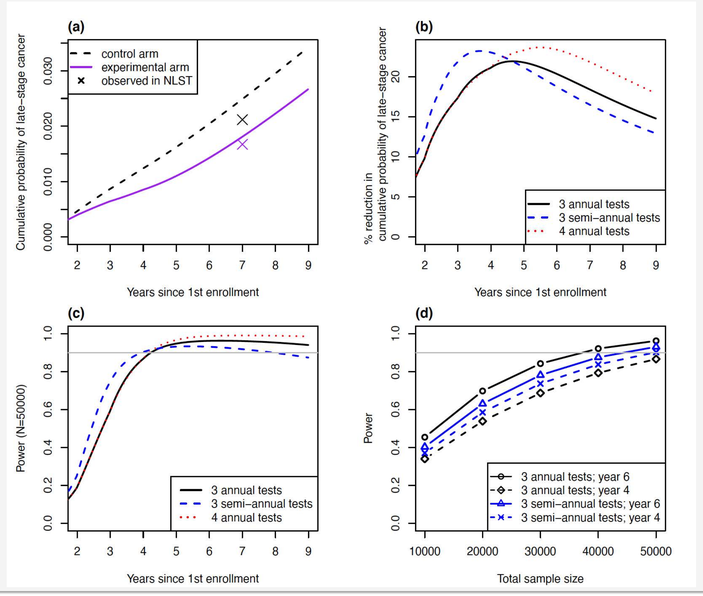
- A 5-state model:



- Early/Late-stage preclinical cancers are asymptomatic. They can be detected via tests with sensitivity β.
- Under the stable disease condition, we specify mean sojourn times via the prevalence of preclinical cancers from a quick pilot study.
- Model-projection of the late-stage cancer incidence also depends on the number and the timing of tests, follow-up time, accrual and adherence.
- With projected incidences in two arms, the statistical powers is calculated approximately via a z-score (standardized difference).

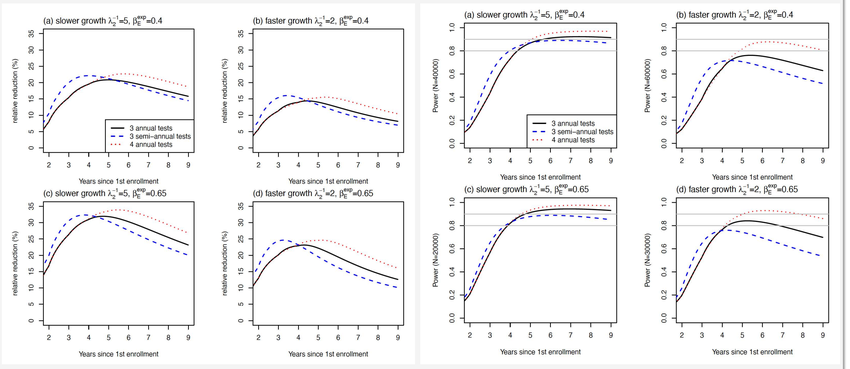
Examples based on the NLST

- The National Lung Screen Trial (2002-2009) randomized eligible smokers to the chest radiography arm (control arm) or the low-dose CT arm (experimental arm).
- The Hu-Zelen model was used for the study design: with a total sample size of 50k, the study with 3-annual tests had statistical power of 90% for detecting a 21% reduction in lung cancer mortality at Yr 6.
- We borrow the same set of hypothesized parameters with additional stage-specific assumptions, including sensitivities:
  - Experimental arm: 85% for late-stage; 51% for early-stage
  - Control arm: 25% for late-stage; 16% for early-stage
- Below figures summarize the results: e.g., 3-annual tests, N=40k, 90% power for a 20% reduction in late-stage incidence at Yr 6.



Additional examples

- Keep the most parameters same as in the NLST but vary two key parameters:
- $\lambda_2^{-1} = (2, 5)$ for faster/slower-growing cancers; narrower/wider window of early-detection
- $\beta^{Early} = (40\%, 65\%)$ for less/more sensitivity tests in the experimental arm



Discussion

- Our model captures the time-varying (delayed & diluted) screening effects / power as results of interaction between timing of tests/analysis and the natural history, which is useful for an optimal study design.
- Our model informs setting of meaningful clinical performance criteria of new biomarker tests during earlier phases of biomarker discovery and validation studies.
- Future work is to extend our model to mortality for studying the surrogacy of endpoints and for suggesting a novel Phase 4/5 study design.

Acknowledgements: *We thank Ruth Etzioni for helpful discussions.*