Projecting surrogate value in the design of cancer screening trials

Arnold Johnsen, Yingqi Zhao, Yingye Zheng

Division of Biostatistics, Fred Hutchinson Cancer Center

Abstract

As cancer screening techniques evolve rapidly, there is increasing demand for quicker evaluation of effectiveness in randomized screening trials while maintaining rigor in estimating mortality benefits. One possible method to accomplish this is using shorter-term outcomes, such as the incidence of late-stage cancer, as a surrogate endpoint for mortality. We propose a general method using microsimulations from a natural history model to evaluate the reliability of such surrogates for a novel screening test and design the trial around the surrogate.

Background: meta-analysis, modeling, pitfalls

- Prior work in establishing surrogacy of late-stage incidence (LSI) for mortality has taken the form of meta-analysis and modeling-based studies.
- Meta-analysis requires many completed trials for a cancer/screening test type, and assumes future tests/trials are comparable to previous ones.
- Many modeling studies [1], [2] assume natural history models where stage is already a surrogate, i.e. all causal pathways between screening and mortality are mediated by stage.
- Work in both fields pursues *causal association*, but important information about the underlying surrogate structure can be illuminated through a *causal effectiveness* paradigm [3].

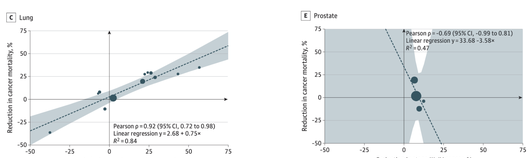

Figure 1: Two panels from Figure 1 of Feng at al. (2024) [4]. The association between LSI and mortality endpoints is strong in lung cancer screening, but weak in prostate cancer.

Research reported in this presentation was funded through a Patient-Centered Outcomes Research Institute® (PCORI®) Award (ME-2023C3-35543). The statements presented in this work are solely the responsibility of the authors and do not necessarily represent the views of the PCORI®.

Causal structures of natural history models

Common factors in the debate over the use of LSI as a surrogate include mode of diagnosis (better prognosis for screen-detected cancers), cancer subtypes (e.g. small-cell lung cancer), and non-stage measures of severity (e.g. tumor size in breast cancer).

These factors can be situated in the causal diagrams below. To assess surrogacy, we propose simulating from a model that either accounts for them directly, or uses a simplified structure to allow for at least one non-stage pathway.

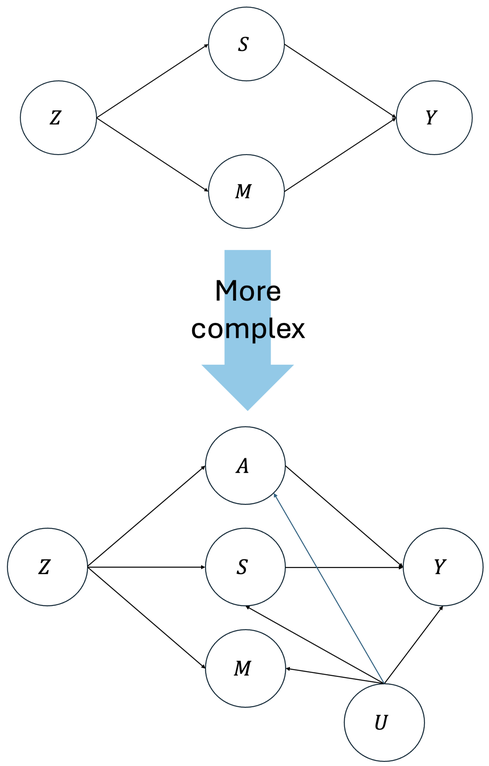

Figure 2: Causal diagrams of cancer screening. Z: treatment arm, A alternative mediator, S LSI, M mode of diagnosis, U confounders, Y mortality.

Importantly, association between the causal effects of Z on Y and Z on S does not completely describe the causal relationships in the diagrams. *Different diagrams can produce the same association measures*, hence the need for a causal effectiveness lens.

Diagnostics for causal association and effectiveness

We use two metrics to assess surrogacy:

1. *Matched projections* simply calculate association between projected endpoints from the microsimulated data.
2. *Causal effect predictiveness (RRCEP)* [5] calculates mortality risk reduction among individuals who did/didn't experience a downstaging benefit due to screening.

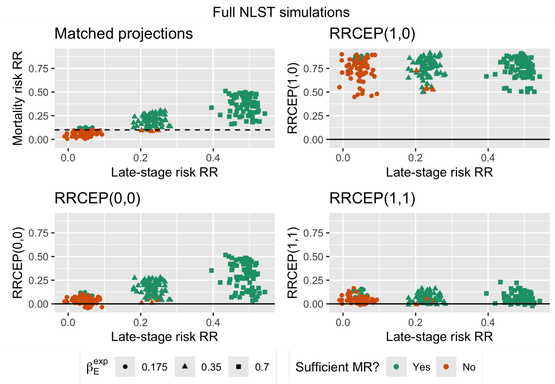

Figure 3: β_E^{exp}: early-stage sensitivity in 5-state model [6].

The top left panel shows that for sufficiently high early-stage sensitivity, a majority of simulated trials hit a prespecified benchmark of 10% mortality reduction. Therefore, a trial that can detect a mortality reduction of around 20% with adequate power should be run.

The remaining panels show that there is a large mortality reduction among downstagers (1,0) but also some within-stage benefit particularly among never-late-stage individuals (0,0). Because the former effects (called *associative*) are larger than the latter (*disassociative*), CEP indicates that stage-based causal pathways dominate non-stage pathways.

Figure 4: Two extremes in causal effectiveness paradigm. In the left panel, $\text{CEP}(0,0) = \text{CEP}(1,1) = 0$. In the right, $\text{CEP}(1,0)$ equals the overall causal effect of Z on Y.

Using surrogate diagnostics

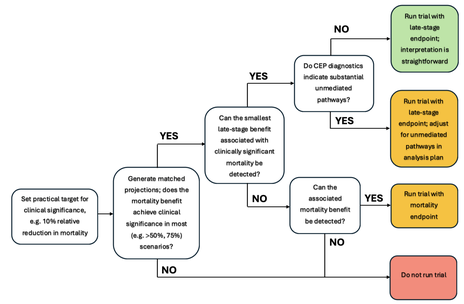

Figure 5: Trial-level decision-making for a proposed surrogate S.

- Matched projections provide go/no-go decisions; if very few trials achieve clinically significant mortality, or the associated surrogate effects cannot be detected with power, do not run the trial.
- CEP measures the relative strength of mediated to unmediated pathways, i.e. differentiates between the two diagrams in Figure 4. Other mediators can be added to S or adjusted for in the prespecified analysis plan.
- The process in Figure 5 can be iterated for different definitions of S, including varying reference times (4-year or 6-year LSI?) to find an optimal surrogate.

References

[1] Lukas Owens, Roman Gulati, and Ruth Etzioni.
Stage shift as an endpoint in cancer screening trials: Implications for evaluating multicancer early detection tests.
Cancer Epidemiology, Biomarkers Prevention, 31(7):1298–1304, 07 2022.

[2] James Y Dai, Jing Zhang, Jerome V Braun, Noah Simon, Earl Hubbell, and Nan Zhang.
Clinical performance and utility: A microsimulation model to inform the design of screening trials for a multi-cancer early detection test.
Journal of Medical Screening, 31(3):140–149, 2024.

[3] Michael R. Elliott.
Surrogate endpoints in clinical trials.
Annual Review of Statistics and Its Application, 10(Volume 10, 2023):75–96, 2023.

[4] Xiaoshuang Feng, Hana Zahed, Justina Onwuka, Matthew E. J. Callister, Mattias Johansson, Ruth Etzioni, and Hilary A. Robbins.
Cancer stage compared with mortality as end points in randomized clinical trials of cancer screening: A systematic review and meta-analysis.
JAMA, 331(22):1910–1917, 06 2024.

[5] Peter B. Gilbert and Michael G. Hudgens.
Evaluating candidate principal surrogate endpoints.
Biometrics, 64(4):1146–1154, 11 2008.

[6] Kehao Zhu, Ying-Qi Zhao, and Yingye Zheng.
Designing cancer screening trials for reduction in late-stage cancer incidence.
Biometrics, 80(3):ujae097, 09 2024.