**Simple optimal cost-effective risk thresholds for a single screen**

Hormuzd A. Katki\*1, Ph.D. and Ionut Bebu2, Ph.D.

1Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, DHHS, Rockville, MD, USA

2Biostatistics Center, George Washington University, Rockville MD, USA

**\*Corresponding author:**

**HAK:** Division of Cancer Epidemiology and Genetics, NCI, 9609 Medical Center Dr., Room 7E592, Rockville, MD 20850, Phone: 240-276-7423, email: [katkih@mail.nih.gov](mailto:katkih@mail.nih.gov)

**Short title:** Cost-effectiveness of a single screen

**Key Words:** incremental net benefit, screening, decision analysis, risk prediction,

**Funding/Support:** This study was supported, in part, by the Intramural Research Program of the US National Institutes of Health/National Cancer Institute.

**Role of the Sponsor:** The NIH had no role in the design and conduct of the study; in the collection, analysis, and interpretation of the data; or in the preparation, review, or approval of the manuscript.

**Conflicts of Interest:**  None

**Acknowledgements:** This research was supported, in part, by the Intramural Research Program of the NIH/National Cancer Institute. We thank William Black (Dartmouth Hitchcock Medical Center) for helping provide cost estimates for aspects of CT lung-cancer screening.

**xx words (max 3000), 6 references, 2 small tables/figures**

**Abstract ( words; max 275)**

**Introduction**:

**Methods**: We demonstrate that the Incremental Net Benefit (INB) is a function of screening test characteristics, life-years gained from screening, and 4 costs (cost of screening test, cost of confirmatory test, cost of treatment for early-detected disease, cost of treatment for late-detected disease).

**Results**:

**Conclusions**:

**Introduction**

A key question in population disease screening is who should be referred for tests that are definitive for the presence of disease. However, definitive tests, such as biopsies or diagnostic imaging, are considered too invasive or costly to use in the absence of symptoms. Thus definitive testing is limited to those who test “positive” on a screening test that is acceptably less costly and invasive than definitive tests. However, screening tests are imperfect indicators of the presence of disease. The ideal goal is choose a screening test threshold that is optimally cost-effective for controlling the rates of false classification.

There is much literature on identifying cost-effective thresholds, such as the classic paper of Pauker and Kassirer. Unfortunately, the cost-effective threshold depends entirely on the costs, benefits, and harms of the tests and subsequent interventions, and for many reasons, these appear to usually be difficult to specify in clinical practice. Subsequent mathematical developments conglomerate costs, benefits and harms into “utilities” that are even harder to specify. Furthermore, metrics that do not require costs, benefits, and harms, such as Youden’s index, AUC, Net Benefit, and Mean Risk Stratification, cannot formally identify cost-effective thresholds.

Because costs, benefits and harms require specification, we propose that scientists who evaluate new medical tests or propose thresholds for test must try to specify these as best as possible. We propose a practical framework for specifying the costs and benefits for a single-time screen, such as for gestational diabetes, syphilis in pregnancy, genetic-mutation screening, or the first screen of a cancer screening program. We show that the Incremental Net Benefit (INB) depends on only 3 quantities: the costs of the screening test and definitive test, and a parameter we call the net effectiveness of early intervention (NE). We show that the optimal threshold depends on only the cost of the definitive test and the NE. Our framework also provides a practical criterion for judging when one screening test is more cost-effective than another.

We apply our framework to two important screening programs. …

**Methods**

*Defining Costs and Effectiveness for a single-time screen*

We define 4 key costs of a single-time screen:

1. *Cm*: The cost of evaluating the marker as the screening test
2. *C0*: The cost of evaluating the definitive test. We presume this test is perfect, but is so costly or invasive that it cannot usually be performed on asymptomatic people.
3. *C1*: The cost of treatment for disease diagnosed early (i.e. disease diagnosed among those with positive screening tests)
4. *C2*: The cost of treatment for disease diagnosed late (i.e. disease occurring either without screening or among those with negative screening tests)

There are 3 key parameters that define the effectiveness of a single screen, defining effectiveness as the value of the gain in life-expectancy:

1. *E0*: The value of the life-expectancy for someone who never develops disease
2. *E1*: The value of the life-expectancy for someone with a positive screening test who is subsequently diagnosed with disease
3. *E2*: The value of the life-expectancy for someone who is either not screened or has a negative screening test, but is subsequently diagnosed with disease.

The “value” must be on the same scale as costs, and this represents the monetary value of a year of life-gained. In our examples of population mutation-screening and CT lung-cancer screening, we derive these 4 cost and effectiveness parameters E1 and E2 (we show that *E0* does not require specification). We propose sensitivity analyses for costs and effectiveness and, in the Discussion, we consider practical weaknesses of our approach to identifying costs and effectiveness.

*Calculating the Incremental Net Benefit (INB)*

The INB considers the difference in costs and effectiveness of never screening versus a 1-time screen. Define disease as *D* and the marker for the screening test as *M*.

1. No screening. There are 2 possibilities:
   1. The person has disease (*D+*): The cost is Cm+C2 and effectiveness is E2.
   2. The person does not have disease (*D*-): Cost is 0 and effectiveness is E0.
2. We screen for disease using marker *M*. The marker is dichotomized as positive and negative based on a underlying threshold that we later discuss how to set. There are 4 possibilities:
   1. The screen is positive and the person has disease (*D+*,*M+*).
      1. Cost is Cm+C0+C1 and effectiveness is E1.
   2. The screen is positive by the person does not have disease .
      1. Cost is and effectiveness is .
   3. The screen is negative but the person develops disease .
      1. Cost is and the effectiveness is .
   4. The screen is negative and the person does not develop disease .
      1. Cost is and effectiveness is .

Let and denote the expected cost without and with screening, respectively:

.

Let and denote the expected effectiveness without and with screening, respectively:

.

Then the increase in cost and effectiveness due to screening versus not screening are

**References**

1. Moyer VA. Screening for lung cancer: US Preventive Services Task Force recommendation statement. *Annals of internal medicine.* 2014;160(5):330-338.

2. Katki HA, Kovalchik SA, Berg CD, Cheung LC, Chaturvedi AK. Development and validation of risk models to select ever-smokers for CT lung cancer screening. *Jama.* 2016;315(21):2300-2311.

3. U.S. Preventive Services Task Force. Draft Research Plan: Lung Cancer: Screening. May 2018; <https://www.uspreventiveservicestaskforce.org/Page/Document/draft-research-plan/lung-cancer-screening1>.

4. Tammemägi MC, Katki HA, Hocking WG, et al. Selection criteria for lung-cancer screening. *New England Journal of Medicine.* 2013;368(8):728-736.

5. Tammemägi MC, Church TR, Hocking WG, et al. Evaluation of the lung cancer risks at which to screen ever-and never-smokers: screening rules applied to the PLCO and NLST cohorts. *PLoS medicine.* 2014;11(12):e1001764.

6. Aberle DR, Adams AM, Berg CD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med.* 2011;365(5):395-409.

Thus much of the work on evaluating screening tests has been based on measures of classification, such as Youden’s index or the AUC, or measures of risk stratification such as the predictiveness curve or Mean Risk Stratification (MRS). The most popular decision-analytic metric, Net Benefit (NB), cannot be used to identify a cost-effective threshold and instead considers thresholds in a sensitivity analysis. Furthermore, Youden’s index, AUC, MRS, and NB are all optimized when the risk threshold equals disease prevalence.