Author Manuscript

Title: Re: Think Before You Leap

Summer S. Han[1,2], Kevin ten Haaf[3], William D. Hazelton[4], Jihyoun Jeon[5], Rafael Meza[5], Chung Yin Kong[6], Eric J. Feuer[7], Harry J. de Koning[3] and Sylvia K. Plevritis[2]

[1]Department of Medicine, Stanford University, Palo Alto, CA

[2]Department of Radiology, Stanford University, Palo Alto, CA

[3]Department of Public Health, Erasmus MC, Rotterdam, The Netherlands

[4]Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA

[5]Department of Epidemiology, University of Michigan, Ann Arbor, MI

[6]Department of Radiology, Massachusetts General Hospital, Boston, MA

[7]Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD

Corresponding Author:

Sylvia K. Plevritis

Department of Medicine

Stanford University School of Medicine

Stanford, CA 94305

Email: Sylvia.plevritis@stanford.edu

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of record. Please cite this article as doi:10.1002/ijc.31183.

This article is protected by copyright. All rights reserved.

Author Manuscript

Dear Editor:

We thank Dr. Detterbeck for his review on our recent study that estimates the overdiagnosis of lung cancer screening and for the opportunity to clarify our work. In our response, we address each of the reviewer's points.

RE: LACK OF ACCOUNTING FOR AN ALTERED SPECTRUM OF AGGRESSIVENESS AMONG SCREEN-DETECTED TUMORS

As Dr. Detterbeck points out, screening can detect non-aggressive tumors not observed in a non-screened cohort, because these indolent tumor would never progress to point of causing symptoms that would prompt detection. Indeed, a significant fraction of the lung cancers (15% of the females and 7.5% of the males in the CT screening arm) in the National Lung Screening Trial (NLST) were identified as bronchioloalveolar carcinoma (BAC) which are regarded as a non-invasive lesion, largely detected on screening. Dr. Detterbeck stated that our study excluded this type of non-invasive indolent lung cancer. He justified this claim by suggesting that our models were derived exclusively from SEER data, thereby representing a non-screening population. In actuality, our models utilize the non-screening population data from SEER as well as data from screening trials. A reader of our manuscript will learn that screening associated parameters of our models were obtained by calibration to NLST and PLCO data [1-4], as stated in our ***Methods***, first paragraph)[1]:

> *Each model was calibrated and validated using the data from NLST and PLCO (Prostate, Lung, Colorectal, and Ovarian Cancer Screening) to obtain estimates on screening- related parameters such as tumor size thresholds for diagnostic follow-up. Each model reproduced the observed incidence and mortality of lung cancer (stratified by cancer stage at diagnosis, histology, sex, and detection mode) in both arms of these trials.*

John Wiley & Sons, Inc.

This article is protected by copyright. All rights reserved.

Author Manuscript

In particular, simulations of the NLST population using the Stanford model show that 15.8% of the female lung cancer cases in the CT screening arm have BAC and 7.7% of the male cases in the same arm have BAC, which is close to the 15% and 7.5% observed in NLST data. **Supplemental Figure 3** in our original paper also shows that the overdiagnosis rate associated with BAC is higher (model range: 10.2%-42.6%) compared to other invasive subtypes such as adenocarcinoma (model range: 3.7%-23.6%) in the population-level simulations. Therefore, Dr. Detterbeck's statement that our models do not account for the type of overdiagnosis resulting from screen detection of indolent tumors is invalid.

RE: Assumptions Regarding Competing Causes of Death

As Dr. Detterbeck points out, another type of overdiagnosis is related to competing causes of death. That is, a patient's lung cancer is considered overdiagnosed if a patient dies of any other cause before the cancer causes clinical symptoms. Dr. Detterbeck states that because our models are based on overall mortality for the US population data of other cause mortality (OCM) rates, these rates would be higher than the OCM rates observed in NLST due to "healthy volunteer effects". Therefore, he claims that our models overstate the type of overdiagnosis due to competing causes of death. We disagree with Dr. Detterbeck on this point. While our models utilize U.S. population mortality data (http://www.mortality.org/) to estimate the rates of mortality from competing causes[5], our simulation results for NLST show that the predicted OCM rates using our models are comparable to the OCM observed in NLST (see **Figure 1**). This is expected because NLST is comprised of heavy smokers with high smoking-related comorbidities, and the OCM predictions in our models are based on individuals' smoking histories, where heavy smokers are likely to have a higher risk of mortality compared to the average U.S. population. For this reason, applying our prediction model for OCM to NLST still produces comparable estimates to the observed data in the trial and therefore remains a valid approach. To be clear, we agree with the premise that competing risks deserves more

John Wiley & Sons, Inc.

This article is protected by copyright. All rights reserved.

consideration, but Dr. Detterbeck's criticisms about our modeling work with respect to OCM are invalid.

RE: Impact of the Modeling Assumptions

As we discuss above, the estimates of overdiagnosis from our models are based on *both* (i) overdiagnosis stemming from screen detecting indolent tumors and (ii) overdiagnosis related to competing causes of death. Our results, based on both of these types of overdiagnosis, show that while screening through age 80 is efficient in reducing lung cancer mortality irrespective of overdiagnosis, stopping screening at a younger age of 75 provides a greater efficiency in reducing lung cancer deaths and increasing life-years per overdiagnosed case.

RE: Using Retrospective Findings to Predict Prospective Outcomes

We agree with Dr. Detterbeck that one must be careful in applying retrospective findings to predict future outcomes because clinical practice patterns and patient behaviors will change over time. However, one can also argue that findings from the past can provide invaluable insights when deciding public health policies for the future, particularly in the case of overdiagnosis. Surprisingly, Dr. Detterbeck states that the concept of overdiagnosis has no direct relevance to a clinician who is advising an individual patient. We disagree. If a certain subgroup of an asymptomatic population is at a higher risk of overdiagnosis, then those individuals should be informed of these potential harms and recommendations can be made to weigh these risk against the benefits associated with screening. In fact, the U.S. Preventive Services Task Force (USPSTF) explicitly mentions that overdiagnosis should be considered in shared-decision making for lung screening (https://www.uspreventiveservicestaskforce.org/Page/Document/RecommendationStatementFinal/lung-cancer-screening):

John Wiley & Sons, Inc.

This article is protected by copyright. All rights reserved.

Author Manuscript

Author Manuscript

"Overdiagnosis of lung cancer and the risks of radiation are real harms, although their magnitude is uncertain. The decision to begin screening should be the result of a thorough discussion of the possible benefits, limitations, and known and uncertain harms."

As Dr. Detterbeck has mentioned, overdiagnosis driven by competing causes of death can be prevented if we avoid screening individuals with limited life expectancy, which may be highly associated with smoking histories and ages in the context of lung screening. To this point, we ask Dr. Detterbeck to consider embracing the modeling paradigm. It provides a platform to test different hypotheses about the future – including changes in clinical risk and practice patterns. However before extrapolating a model to under-studied areas, one must make sure the model is consistent with all the available data ("the past"). Herein lies the significance of the work that we have provided: a rigorous model-based analysis that reproduces the wealth of the past information (coming from lung screening trials, population registry data, and prospective cohort studies) is used to predict the future impact of alternative screening implementation scenarios.

Re: Think Before You Leap

While we welcome a robust discussion on the scientific merits of our work and the relevance of modeling to clinical decision-making, we do take exception on Dr. Detterbeck's main criticism "Think before your leap." We want to reassure Dr. Detterbeck and our readers that we did not leap to any conclusions. We proceeded in a methodologically rigorous manner. To mitigate the uncertainty associated with model building, we presented a comparative modeling analysis (four models in this analysis). In this response, we argued that all of Dr. Detterbeck's criticisms on the scientific merits of our work are invalid. However we do agree with the underlying spirit of Dr. Detterbeck's comments, which seems to be: what should we do next? Let's employ modeling in the deliberations of complex issues ahead to heed the words "Think Before You Leap."

John Wiley & Sons, Inc.

This article is protected by copyright. All rights reserved.

Author Manuscript

Acknowledgements

This report is based on research conducted by the National Cancer Institute's (NCI) Cancer Intervention and Surveillance Modeling Network (CISNET) (NIH grants U01-CA152956 & U01-CA199284). K.T.H. is affiliated with the Dutch-Belgian NELSON trial. He has received grants from: 1) Sunnybrook health sciences, Toronto, Canada to evaluate the cost-effectiveness of lung cancer screening in Ontario and 2) the university of Zurich to evaluate the cost-effectiveness of lung cancer screening in Switzerland. He is an invited speaker at the 4th IASLC SSAC CT screening workshop on 12/3/2016; travel expenses were in part paid by the IASLC SSAC organizing committee.

References

1. Han SS, ten Haaf K, Hazelton WD, Munshi VN, Jeon J, Erdogan SA, Johanson C, McMahon PM, Meza R, Kong CY. The impact of overdiagnosis on the selection of efficient lung cancer screening strategies. International Journal of Cancer 2017.

2. Han SS, Erdogan SA, Toumazis I, Leung A, Plevritis SK. Evaluating the impact of varied compliance to lung cancer screening recommendations using a microsimulation model. Cancer Causes & Control 2017;28:947-58.

3. de Koning HJ, Meza R, Plevritis SK, Ten Haaf K, Munshi VN, Jeon J, Erdogan SA, Kong CY, Han SS, van Rosmalen J. Benefits and harms of computed tomography lung cancer screening strategies: a comparative modeling study for the US Preventive Services Task Force. Annals of internal medicine 2014;160:311-20.

4. Meza R, Haaf K, Kong CY, Erdogan A, Black WC, Tammemagi MC, Choi SE, Jeon J, Han SS, Munshi V. Comparative analysis of 5 lung cancer natural history and screening models that reproduce outcomes of the NLST and PLCO trials. Cancer 2014;120:1713-24.

John Wiley & Sons, Inc.

This article is protected by copyright. All rights reserved.

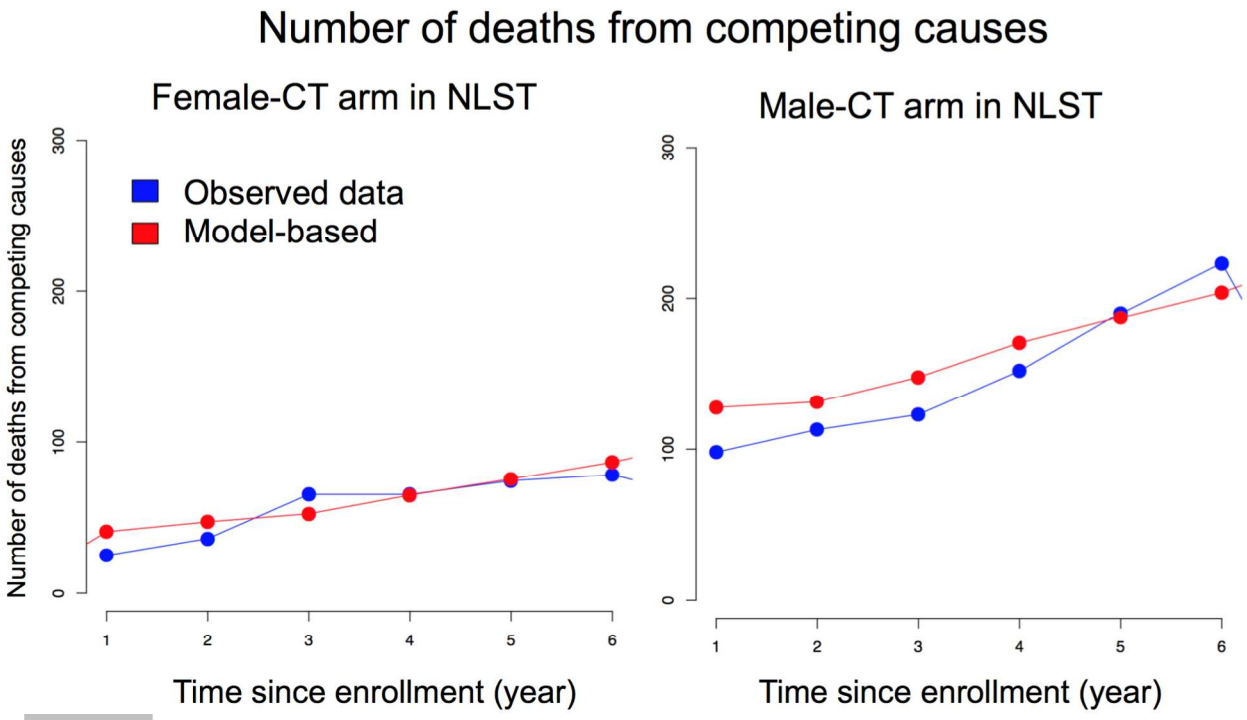
Author Manuscript

5. Rosenberg MA, Feuer EJ, Yu B, Sun J, Henley SJ, Shanks TG, Anderson CM, McMahon PM, Thun MJ, Burns DM. Cohort life tables by smoking status, removing lung cancer as a cause of death. Risk Analysis 2012;32.

John Wiley & Sons, Inc.

This article is protected by copyright. All rights reserved.

Figure 1. Comparing the number of deaths from competing causes in NLST CT screening arm: observed vs. model based using Stanford model



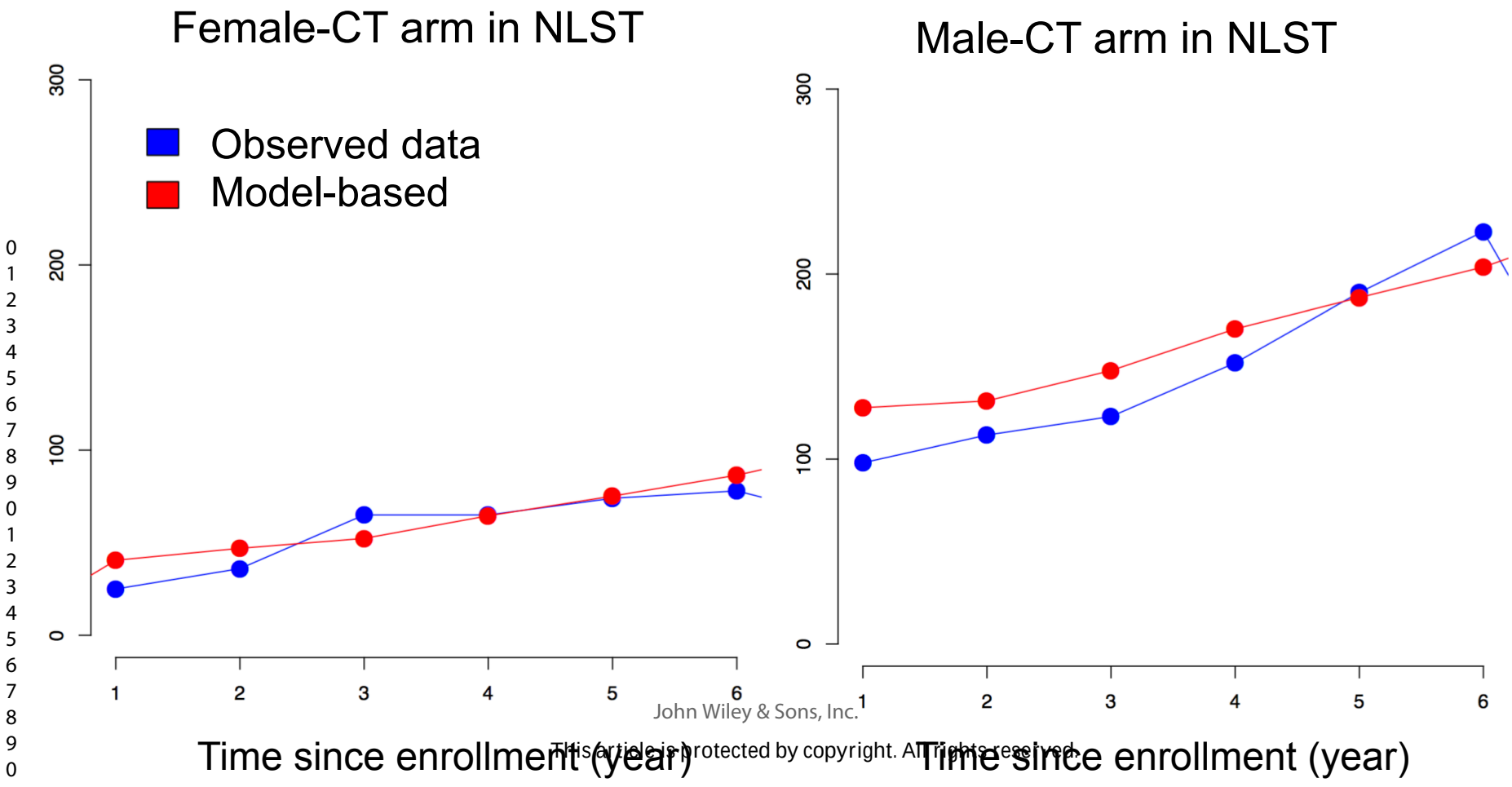
Author M

John Wiley & Sons, Inc.

This article is protected by copyright. All rights reserved.

International Journal of Cancer

Number of deaths from competing causes



John Wiley & Sons, Inc.
This article is protected by copyright. All rights reserved.