Title
RE: A predictive model for lung cancer screening nonadherence in a community setting healthcare network

Journal
JNCI Cancer Spectrum, 8(3)

ISSN
2515-5091

Authors
Lin, Yannan
Ding, Ruiwen
Petousis, Panayiotis
et al.

Publication Date
2024-04-30

DOI
10.1093/jncics/pkae040

JNCI Cancer Spectrum, 2024, **8(3),** pkae040
https://doi.org/10.1093/jncics/pkae040
Advance Access Publication Date: May 23, 2024
Correspondence

RE: A predictive model for lung cancer screening nonadherence in a community setting healthcare network

Yannan Lin, MD, MPH, PhD,[1,2,*] Ruiwen Ding, BS,[1,2] Panayiotis Petousis, PhD,[3] Ashley Elizabeth Prosper, MD,[2] Denise R. Aberle, MD,[1,2] William Hsu, PhD[1,2]

[1]Department of Bioengineering, University of California, Los Angeles, CA, USA
[2]Medical & Imaging Informatics, Department of Radiological Sciences, David Geffen School of Medicine at the University of California, Los Angeles, Los Angeles, CA, USA
[3]UCLA Clinical and Translational Science Institute, David Geffen School of Medicine at the University of California, Los Angeles, Los Angeles, CA, USA
*__Correspondence to:__ Yannan Lin, MD, MPH, PhD, 924 Westwood Blvd Ste 420, Room K, Los Angeles, CA 90024, USA (e-mail: yannanlin@mednet.ucla.edu).

To the Editor:

We acknowledge the invaluable contributions of Bastani et al. (1) in their recent publication on predicting patient nonadherence to the follow-up recommendations made at the time of baseline (first) lung cancer screening (LCS) exams. The authors developed machine-learning-based models using 9 clinical and demographic variables in a retrospective cohort of 1875 patients. The best-performing gradient-boosting model achieved a cross-validated area under the receiver operating characteristics curve (AUC) of 0.89 (95% confidence interval [CI] = 0.87 to 0.90).

Given the importance of ensuring adherence to screening outcomes (2,3), we sought to implement their model, parameterizing it using data from our institution. Unfortunately, we were unable to attain the same level of predictive performance using these 9 predictors to predict nonadherence to baseline LCS recommendations in our cohort. A total of 2430 eligible patients who underwent a baseline low-dose computed tomography (LDCT) screen at our institution between July 31, 2013, and November 30, 2021, were included. We used the same definition of nonadherence as Bastani et al. and retained the same categories for the predictors. Random forest yielded the highest cross-validated AUC (10-fold with grid search) of 0.68 (95% CI = 0.66 to 0.69), considerably lower than the reported 0.89 by Bastani et al. Our previous analysis using 6 clinical, demographic, and health-related variables yielded a similar test AUC (4). The data underlying this study cannot be shared publicly to protect the privacy of study participants.

Low adherence in clinical LCS programs is concerning (5). Models capable of accurately predicting nonadherence to LCS recommendations are pivotal in identifying patients at a high risk of nonadherence, allowing targeted interventions with limited clinical resources. One potential reason existing adherence prediction models do not generalize is due to underspecification (ie, underfitting due to a lack of key predictors). Many of the variables that we suspect may improve model performance are not routinely readily available in the electronic medical record, and further research may focus on additional patient-level [eg, smoking-related stigma (6)], physician-level (eg, prior LCS experience), and system-level factors associated with LCS nonadherence. Another potential reason the identified variables did not have the same predictive value in our cohort is a difference in the target population (see Table 1). Leveraging multicenter datasets may improve our ability to identify robust predictors against distribution shifts across institutions. Additionally, we observed changes in patient characteristics within our LCS programs, such as a reduction in pack-years, after the release of the 2021 United States Preventive Services Task Force LCS guidelines. Since most patients from the two studies were included before implementing the 2021 guidelines, refining models to account for this change using prospective data may prove essential. Finally, our results are based on our implementation of the authors' model. Making their prediction model open-source would help conduct external validation studies and facilitate future adoption.

While we applaud the work of Bastani et al. in building a prediction model for nonadherence, we hope to emphasize the need for and encourage additional work within the scientific community to improve model specificity and generalizability for effective clinical implementation.

Data availability

The study was performed with institutional review board approval and waiver of informed consent, and the data underlying this study cannot be shared publicly due to the privacy of individuals who participated in the study. Aggregated summaries without individual data were shared in the article.

Author contributions

Yannan Lin, MD, MPH, PhD (Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Project administration; Validation; Writing—original draft; Writing—review & editing), Ruiwen Ding, BS (Conceptualization; Investigation; Methodology; Writing—review & editing), Panayiotis Petousis,

Received: April 12, 2024. **Accepted:** May 16, 2024

Table 1. Comparison of patient characteristics at the baseline screening between the two studies[a]

Variable	Individual, No. (%) UCLA (N = 2430)	Bastani et al. (N = 1875)	P[d]
Sex			.002
Female	977 (40.2)	845 (45.1)	
Male	1453 (59.8)	1030 (54.9)	
Age in years			<.001
55-59	510 (21.0)	562 (30.0)	
60-69	1244 (51.2)	983 (52.4)	
≥70	676 (27.8)	330 (17.6)	
Race			<.001
Asian	203 (8.4)	100 (5.3)	
Black	171 (7.0)	170 (9.1)	
White	1933 (79.5)	1501 (80.1)	
Other[a]	59 (2.4)	0 (0)	
Unknown	64 (2.6)	104 (5.5)	
Lung-RADS			<.001
1	362 (14.9)	531 (28.3)	
2	1677 (69.0)	1081 (57.7)	
3	176 (7.2)	152 (8.1)	
4A	116 (4.8)	72 (3.8)	
4B	72 (3.0)	23 (1.2)	
4X	27 (1.1)	16 (0.9)	
Smoking status			<.001
Current	955 (39.3)	998 (53.2)	
Former	1443 (59.4)	856 (45.7)	
Unknown	32 (1.3)	21 (1.1)	
Site			Not comparable
1	848 (34.9)	152 (8.1)	
2	257 (10.6)	994 (53.0)	
3	62 (2.6)	729 (38.9)	
4	847 (34.9)	0 (0)	
5	2 (0.1)	0 (0)	
6	390 (16.0)	0 (0)	
7	24 (1.0)	0 (0)	
Median household income			<.001
<85k/y	1747 (71.9)	904 (48.2)	
85-100k/y	243 (10.0)	484 (25.8)	
>100k/y	398 (16.4)	487 (26.0)	
Unknown	42 (1.7)	0 (0)	
Referral specialty			<.001
Internal or family medicine	1944 (80.0)	881 (47.0)	
Pulmonary	405 (16.7)	715 (38.1)	
Thoracic	14 (0.6)	112 (6.0)	
Physician assistant or nurse practitioner	0 (0)	73 (3.9)	
Other[b]	67 (2.8)	94 (5.0)	
Insurance			<.001
Medicaid	21 (0.9)	218 (11.6)	
Medicare	1018 (41.9)	777 (41.4)	
Private	1347 (55.4)	774 (41.3)	
Other[c]	39 (1.6)	106 (5.7)	
Unknown	5 (0.2)	0 (0)	

[a] Subcategories in other race: American Indian or Alaska Native, Native Hawaiian or Pacific Islander, more than one race, or other racial groups not otherwise stated. Lung-RADS = Lung Computed Tomography Screening Reporting & Data System; UCLA = University of California, Los Angeles; VA = Veterans Administration.
[b] Subcategories in other referring physician types: other specialties not specified above.
[c] Subcategories in other insurance: Veterans Administration, self-pay, and other insurance not specified.
[d] The P-values are from two-sided χ^2 or Fisher's exact tests.

PhD (Conceptualization; Investigation; Methodology; Writing—review & editing), Denise R. Aberle, MD (Conceptualization; Funding acquisition; Investigation; Methodology; Resources; Supervision; Validation; Writing—review & editing), Ashley Elizabeth Prosper, MD (Conceptualization; Investigation; Methodology; Resources; Supervision; Validation; Writing—review & editing), William Hsu, PhD (Conceptualization; Funding acquisition; Investigation; Methodology; Project administration; Resources; Supervision; Validation; Writing—review & editing).

Funding

This study was supported by funding from the V Foundation (Dr Hsu, Ms Ding) and by grants R01 CA210360 (Drs. Lin, Aberle, and Hsu) and R01CA226079 (Drs. Aberle and Hsu) from the National Institutes of Health (NIH); 1722516 from the National Science Foundation (Dr Hsu); NIH/National Cancer Institute U2C CA271898 (Drs. Aberle, Hsu, and Prosper). The NIH National Center for Advancing Translational Science (NCATS) University of California, Los Angeles (UCLA) Clinical and Translational Research Center (CTSI) Grant Number TL1TR001883, along with the Integrated Diagnostics (IDx) Shared Resource (a joint initiative between the Departments of Radiological Sciences and Pathology and Laboratory Medicine, David Geffen School of Medicine at UCLA), facilitated data access.

Conflicts of interests

Yannan Lin reports financial support was provided by the National Institutes of Health (NIH). Ruiwen Ding reports financial support was provided by the V Foundation. Ashley Elizabeth Prosper reports financial support was provided by the NIH National Cancer Institute. Denise R. Aberle reports financial support was provided by NIH. Denise R. Aberle reports financial support was provided by NIH National Cancer Institute. William Hsu reports financial support was provided by the V Foundation. William Hsu reports financial support was provided by the NIH National Institute of Biomedical Imaging and Bioengineering. William Hsu reports financial support was provided by the NIH National Cancer Institute. William Hsu reports financial support was provided by the National Science Foundation. Denise R. Aberle reports a relationship with the Kaiser Foundation Research Institute (Patient-Centered Outcomes Research Institute) that includes funding grants. Denise R. Aberle reports a relationship with DECAMP 1 PLUS (Janssen Pharmaceuticals) that includes funding grants. Denise R. Aberle reports a relationship with V Foundation that includes funding grants. Denise R. Aberle reports a relationship with LungLife AI that includes funding grants. Denise R. Aberle reports a relationship with Liquid Diagnostics LL that includes funding grants. Denise R. Aberle reports a relationship with EarlyDiagnostics that includes funding grants. William Hsu reports a relationship with EarlyDiagnostics that includes funding grants. William Hsu reports a relationship with the Radiological Society of North America that includes consulting or advisory.

References

1. Bastani M, Chiuzan C, Silvestri G, et al. A predictive model for lung cancer screening nonadherence in a community setting health-care network. *JNCI Cancer Spectr*. 2023;7(2):pkad019. doi: 10.1093/jncics/pkad019
2. Aberle DR, Adams AM, Berg CD, et al.; National Lung Screening Trial Research Team. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med*. 2011;365(5): 395-409. doi:10.1056/NEJMoa1102873

3. de Koning HJ, van der Aalst CM, de Jong PA, et al. Reduced lung-cancer mortality with volume CT screening in a randomized trial. *N Engl J Med*. 2020;382(6):503-513. doi:10.1056/NEJMoa1911793
4. Lin Y, Liang LJ, Ding R, Prosper AE, Aberle DR, Hsu W. Factors associated with nonadherence to lung cancer screening across multiple screening time points. *JAMA Netw Open*. 2023;6(5):e2315250. doi:10.1001/jamanetworkopen.2023.15250
5. Lin Y, Fu M, Ding R, et al. Patient adherence to lung CT screening reporting & data system-recommended screening intervals in the United States: a systematic review and meta-analysis. *J Thorac Oncol*. 2022;17(1):38-55. doi:10.1016/j.jtho.2021.09.013
6. Carter-Harris L, Ceppa DP, Hanna N, Rawl SM. Lung cancer screening: what do long-term smokers know and believe? *Health Expect*. 2017;20(1):59-68. doi:10.1111/hex.12433

JNCI Cancer Spectrum, 2024, 8, 1–3
https://doi.org/10.1093/jncics/pkae040
Correspondence