Supporting Information A: Supplementary Methods

Treatment Capacity Required for Full-Scale Implementation of Lung Cancer Screening in the United States

Erik F. Blom, MD; Kevin ten Haaf, PhD; Douglas A. Arenberg, MD; Harry J. de Koning, MD PhD

Analysis of National Cancer Database data

Data

Lung cancer treatment patterns were derived from the National Cancer Database (NCDB) Participant User File for lung cancer cases diagnosed in the United States between years 2010-2014.

Case selection

We selected cases with International Classification of Diseases for Oncology 3rd edition lung cancer topography codes (C340 - C343, C348 and C349) and malignant behavior code.[1] We included cases that were staged using the American Joint Committee on Cancer 7th edition Cancer Staging Manual, which was effective from 2010 through 2017.[2] Reasons for exclusion were: (1) stage 0, occult stage, unknown stage, or unknown stage subcategory (e.g. stage I rather than stage IA); (2) more than 4 months (122 days) between date of diagnosis and onset of therapy; (3) patient did not receive any treatment at the reporting facility; (4) date of diagnosis before facility's reference date (i.e. the date from which the facility guarantees the accuracy of data); (5) cases with a history of multiple primary tumors of which lung cancer wasn't the first; (6) cases with unknown treatment; and (7) patients aged younger than 35 at diagnosis.
This resulted in a final selection of 440,566 cases.

Derivation of variables

In order to match NCDB data with the MISCAN-lung model, we obtained gender, age, stage and histology specific proportions of patients receiving surgery, radiotherapy, chemotherapy, or no therapy as first course treatment. The derivation of these variables is detailed below.

Derivation of age

Age at diagnosis was used as a continuous variable for ages 35-89. However, the NCDB aggregates data for ages 90 and over.

Derivation of stage

We used clinical stage at diagnosis, as defined by the American Joint Committee on Cancer 7th edition Cancer Staging manual.[2] We did not include stage 0, occult, and unknown cancers because these do not have a clear standard treatment. Also, the MISCAN-lung model does not include these stages. We collapsed stage at diagnosis into the following categories to match MISCAN-lung output: IA, IB, II, IIIA, IIIB, IV.

Derivation of histology

We classified International Classification of Diseases for Oncology 3rd Edition morphological codes into MISCAN-lung histology categories *adenocarcinoma* (including *bronchioalveolar carcinoma* and *large cell carcinoma*)*, squamous cell carcinoma, other non-small cell lung cancer* and *small cell lung cancer*. This classification was based on an earlier publication.[3]

Derivation of treatment variables

The NCDB records the first course of treatment, defined as all methods of treatment recorded in the treatment plan and administered to the patient before disease progression

or recurrence. The NCDB includes treatment modalities *surgery, radiotherapy, chemotherapy, immunotherapy, hormone therapy,* and *other therapy*. If a patient received any of these therapies, they were coded as having received *any therapy.* For this current study, we were not able to separately report on the use of *hormone therapy immunotherapy,* and *other therapy* because these were recorded very infrequently. Patients could receive multiple treatments. Hence, treatment proportions for *surgery, radiotherapy,* and *chemotherapy* do not add up to those for *any therapy*. We were not able to distinguish whether multiple therapies were given concurrently or sequentially.

References

1. Fritz A, Percy C, Jack A, et al. *International Classification of Diseases for Oncology*. Geneva, Switzerland: World Health Organization; 2013.
2. Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. *Ann Surg Oncol.* 2010;17:1471-1474.
3. Ten Haaf K, van Rosmalen J, de Koning HJ. Lung cancer detectability by test, histology, stage, and gender: estimates from the NLST and the PLCO trials. *Cancer Epidemiol Biomarkers Prev.* 2015;24:154-161.