



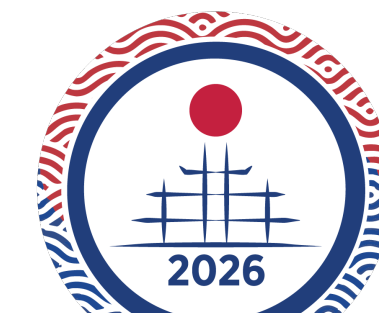
Saw Swee Hock
School of Public Health

Development and Validation of Lung Cancer Risk Prediction Models in the Singapore Chinese Population

Yah Ru Juang¹, Wenjia Chen¹, Alvin Jing Hui Ng¹, Xueling Sim¹, Darren Wan-Teck Lim^{2,3}, Jian-Min Yuan^{4,5}, Adeline Seow¹, Wei Jie Seow^{1,6}

¹Saw Swee Hock School of Public Health, National University of Singapore, Singapore; ²Division of Medical Oncology, National Cancer Centre Singapore, Singapore; ³Duke-NUS Medical School, National University of Singapore, Singapore; ⁴Hillman Cancer Centre, University of Pittsburgh Medical Center, Pittsburgh, PA, USA; ⁵Department of Epidemiology, School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA; ⁶Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore and National University Health System, Singapore

Correspondence: Wei Jie Seow; ephswj@nus.edu.sg



IASLC 2026 World
Conference on Lung Cancer

SEPTEMBER 12 – 15, 2026
SEOUL, REPUBLIC OF KOREA

INTRODUCTION & AIMS

- Lung cancer is the leading cause of cancer death globally.¹
- 5-year survival is only **10–20%**, as most cases are diagnosed at advanced stages.²

Never-Smoker Gap

- Current LDCT screening criteria target heavy smokers (50–80 years; ≥20 pack-years).³
- In Singapore, most lung cancer cases, especially among women, occur in never-smokers, who are not eligible for screening.⁴

Limitations of Existing Risk Prediction Models

- PLCO_{m2012} performs well in Western settings but tends to underestimate lung cancer risk in Asians, particularly low-intensity smokers and long-term quitters.⁵
- Most Asian risk prediction models lack external validation, limiting generalisability.⁶

Study Aim:

- Develop and externally validate **10-year lung cancer risk prediction models** for the Singapore Chinese population.
- Build **sex- and smoking-specific subgroup models** to support personalised risk stratification.

METHODS

Study population (Figure 1)

- Development cohort:** Singapore Chinese Health Study (SCHS)
 - N=63,187 (Chinese adults, aged 45–74 years)
 - Recruitment period (1993 – 1998)
 - Median follow-up years (28.0 years)
 - 813 lung cancer cases over a 10-year follow-up
- Validation cohort:** Multi-Ethnic Cohort (MEC1)
 - N=12,944 (adults, aged ≥18 years)
 - Recruitment period (2004–2010)
 - Median follow-up years (15.4 years)
 - 48 lung cancer cases over a 10-year follow-up

Analytical approach

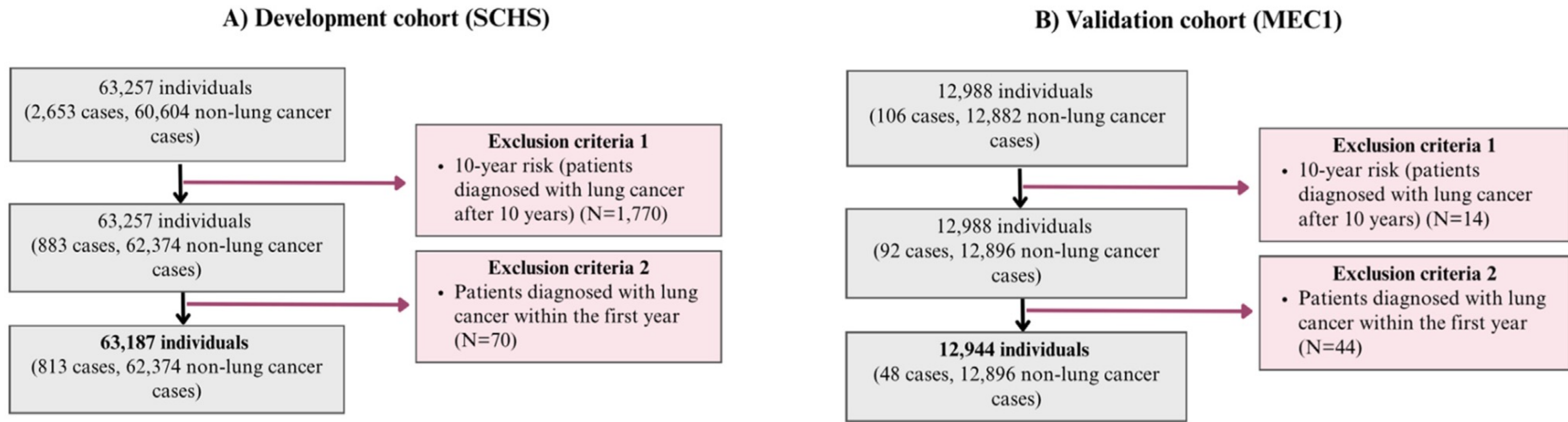
- 29 candidate predictors:** Demographics, smoking history, environmental, reproductive & dietary factors
- Missing data:** MissForest random forest imputation
- Feature selection:** Elastic Net Cox regularisation with 10-fold cross-validation and backward stepwise selection based on AIC
- Internal validation:** 200 bootstrap resamples
- External validation** in Chinese MEC1 participants + sensitivity analyses in Malays & Indians

Model performance

- For overall model, sex-, and smoking-specific models**
 - Discrimination:** Harrell's C-index
 - Calibration:** Calibration intercept and slope
 - Sensitivity & specificity** versus USPSTF criteria

RESULTS

Figure 1. Participant flowchart for the A) development (SCHS) and B) validation (MEC1) cohort



Baseline characteristics of development cohort (Lung cancer vs non-lung cancer participants)

- Compared with non-lung cancer patients, cases were significantly **older** (median 64 vs. 56 years), more likely to be **male** (68% vs. 44%), and **ever-smokers** (70% vs. 30%)
- Cases had significantly **higher pack-years** (median 39.4 vs. 21.9), **lower education** (84% vs. 72% with ≤6 years schooling), and **greater ETS exposure** (87% vs. 82%)

Figure 2. Summary of the multivariable Cox regression model for overall population, sex, and smoking subgroups

Covariate	Overall Model (N=63,187)	Males Model (N=27,897)	Females Model (N=35,290)	Never-Smokers (N=43,918)	Ever-Smokers (N=19,269)
Age (years)	HR 1.09	HR 1.09	HR 1.08	HR 1.09	HR 1.08
BMI (kg/m ²)	HR 0.95	HR 0.95	HR 0.96	—	HR 0.94
Education ≤6 yrs	HR 1.30	HR 1.56	—	—	HR 1.64
Sex (Male)	HR 1.27	—	—	HR 1.32	—
Ever-Smoker	HR 2.14	HR 1.92	HR 2.27	—	—
Pack-years	HR 1.02	HR 1.02	HR 1.02	—	HR 1.02
ETS Exposure	HR 1.21	—	HR 1.55	HR 1.26	—
Cancer History	—	—	—	HR 2.02	—
Allergic Rhinitis	—	—	HR 2.35	HR 2.12	—
Asthma	—	—	HR 1.31	—	—
Fruit Consumption	HR 0.85	HR 0.83	—	—	HR 0.81
Tea Consumption	HR 0.85	—	HR 0.71	—	—
Having Children	—	—	HR 0.53	—	—
Model Performance	C-index: 0.821	C-index: 0.785	C-index: 0.769	C-index: 0.702	C-index: 0.734

■ Risk factor (HR >1) ■ Protective factor (HR <1) ■ Not included in model

- Age (HR: 1.08–1.09) and ever-smoking (HR: 1.92–2.27) were top risk factors across all models
- Higher BMI was consistently protective across all subgroups (HR: 0.94–0.96)
- Never-smokers:** ETS (HR: 1.26), allergic rhinitis (HR: 2.12), and personal cancer history (HR: 2.02) drove lung cancer risk
- Females:** Allergic rhinitis (HR: 2.35), ETS (HR: 1.55), and having children (HR: 0.53, protective) were unique predictors
- Ever-smokers & Males:** Smoking history remains the top risk factor, followed by low education (HR: 1.56–1.64)

Table 1: Model performance in the overall population, sex-specific, and smoking-specific subgroups

	C-index (95 % CI)	Calibration intercept (95% CI)	Calibration slope (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Overall model	0.800 (0.785, 0.814)	0.134 (-0.123, 0.329)	1.012 (0.943, 1.058)	0.778 (0.727, 0.827)	0.681 (0.628, 0.728)
Females model	0.769 (0.743, 0.800)	0.121 (-0.321, 0.644)	1.004 (0.914, 1.102)	0.639 (0.457, 0.817)	0.739 (0.578, 0.875)
Males model	0.785 (0.769, 0.801)	0.298 (-0.043, 0.581)	1.014 (0.934, 1.082)	0.754 (0.680, 0.812)	0.705 (0.66, 0.759)
Ever-smokers model	0.734 (0.713, 0.752)	0.176 (-0.142, 0.446)	1.008 (0.932, 1.089)	0.823 (0.795, 0.853)	0.545 (0.520, 0.563)
Never-smokers model	0.702 (0.670, 0.734)	0.104 (-0.688, 0.752)	0.986 (0.842, 1.123)	0.797 (0.680, 0.885)	0.494 (0.364, 0.614)

External validation of overall model demonstrated good discrimination

- Chinese MEC: **C-index = 0.821** (95% CI: 0.739–0.880)
- Malay/Indian cohort: **C-index = 0.818** (95% CI: 0.702–0.903)

Well calibrated in external validation

- Calibration slopes close to **1.0** (0.984–1.023)
- Moderate calibration-in-the-large (0.290–0.323)

Overall model outperforms USPSTF criteria

- Sensitivity:** 77.8% (95% CI: 72.7–82.7%) vs 53.9% (USPSTF)
- Specificity:** 68.1% (95% CI: 62.8–72.8%) vs 85.3% (USPSTF)

Key finding: Improved lung cancer detection, capturing high-risk individuals missed by current guidelines.

CONCLUSIONS

- First externally validated 10-year lung cancer risk model for the overall Singapore Chinese population using two large prospective cohorts.
- Epidemiological models incorporated region-specific predictors (ETS exposure, dietary factors) are practical for clinical use without genetic testing.
- Sex- and smoking-specific subgroup models revealed distinct risk profiles, offering additional insights beyond smoking history, and with further real-world evaluation, may support more tailored risk stratification in Singapore's clinical settings.**

REFERENCES

- World Health Organization. Global Cancer Burden Growing, Amidst Mounting Need for Services; 2024 Available from: <https://www.who.int/news/item/01-02-2024-global-cancer-burden-growing-amidst-mounting-need-for-services>
- Lu T, Yang X, Huang Y, et al. Trends in the incidence, treatment, and survival of patients with lung cancer in the last four decades. *Cancer Manag Res.* 2019;11:943–953.
- US Preventive Services Task Force, Krist AH, Davidson KW, et al. Screening for Lung Cancer: US Preventive Services Task Force Recommendation Statement. *JAMA.* 2021;325(10):962-970. doi:10.1001/jama.2021.1117
- Juang YR, Hu M, Hui M, et al. Trends in lung cancer incidence and stage at diagnosis in Singapore: a population-based joinpoint regression analysis by age, sex, and smoking status. *Lancet Reg Health West Pac.* 2025;66:101779. Published 2025 Dec 19. doi:10.1016/j.lanwpc.2025.101779
- Yang JJ, Wen W, Zahed H, et al. Lung Cancer Risk Prediction Models for Asian Ever-Smokers. *J Thorac Oncol.* 2024;19(3):451-464. doi:10.1016/j.jtho.2023.11.002
- Juang YR, Ang L, Seow WJ. Predictive performance of risk prediction models for lung cancer incidence in Western and Asian countries: a systematic review and meta-analysis. *Sci Rep.* 2025;15(1):4259. Published 2025 Mar 4. doi:10.1038/s41598-024-83875-6

ABBREVIATIONS

BMI: Body mass index; CI: Confidence interval; C-index: Concordance index; ETS: environmental tobacco smoke; HR: Hazard ratio; LDCT: Low-dose computed tomography; MEC1: Multi-Ethnic Cohort; PLCOm2012: Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial modified 2012 risk prediction model; SCHS: Singapore Chinese Health Study; USPSTF: U.S. Preventive Services Task Force

DISCLOSURE & ACKNOWLEDGEMENTS

Presenter's conflict of interest disclosure: Juang Yah Ru reports no conflict of interest. **AI Use Disclosure:** No AI tools were used in the development of this material.

Acknowledgements: We thank all study participants, investigators, and the research team. We would also like to thank Professor Koh Woon Puay and the National Registry of Diseases Office (NRDO) for their support. This study was supported by the US NIH, Singapore NMRC, BMRC, Ministry of Health (PHMA), National University of Singapore, and National University Health System.