Lung Cancer Risk Prediction Models for Asian Ever-Smokers

Jae Jeong (JJ) Yang, PhD, MPH

University of Florida, Gainesville, USA

September 26, 2024

Nothing to disclose

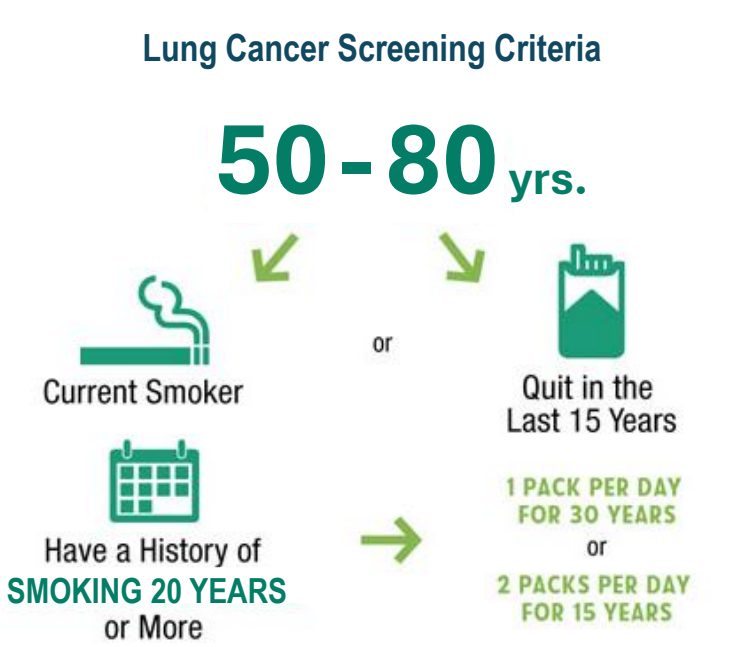
WCE

WORLD CONGRESS OF EPIDEMIOLOGY 2024

Lung Cancer Screening

Implementing Low-Dose Computed Tomography Scan for Early Detection

US Preventive Services Task Force (USPSTF)
Annual Screening for Lung Cancer with
Low-dose computed tomography (LDCT)



The US Preventive Services Task Force (USPSTF) recommendation 2020

- **Benefits**
  : Early detection at the curable stage
  : Lower the chances of dying from lung cancer among high-risk individuals (smokers)

- **Concerns**
  : Uncertainty on the benefit-to-harm ratio
  : Possibility of false-positive results leading to unnecessary invasive procedures/complications

- How to ***better select the LDCT screening candidates*** who would benefit most?

- **Personalized lung cancer risk assessment**
  : To determine screening eligibility incorporating a more comprehensive smoking history and other potential risk factors

Lung Cancer Risk Prediction Models

Selection of Ever-Smokers for Computed Tomography Lung Cancer Screening

Model	Targeted Outcome	Time Frame	Risk Factors Incorporated in Models
Bach (2003)	LC Incidence	1-10 y	age, gender, smoking duration, cigarettes smoked per day, years since cessation, asbestos exposure
Spitz (2007)	LC Incidence	1 y	age, gender, history of COPD, family history of any cancer (2 or more), family history of smoking related-cancer (1 or more), age at quitting, pack-years, asbestos exposure, dust exposure, history of hay fever
LLP V1, V2, V3 (2008)	LC Incidence	5 y	age, gender, smoking duration, history of cancer and pneumonia, family history of lung cancer (early/late onset), asbestos exposure
Hoggart (2012)	LC Incidence	1 y	smoking status, smoking duration, age at staring smoking, cigarettes smoked per day
$PLCO_{M2012}$ (2013)	LC Incidence	6 y	age, race, education, BMI, history of cancer and COPD, family history of lung cancer, smoking status, smoking duration, cigarettes smoked per day, years since cessation
LLPi (2015)	LC Incidence	8.7 y	age, gender, smoking duration, history of cancer and COPD, family history of lung cancer (early/late onset)
Pittsburgh (2015)	LC Incidence	6 y	age, smoking status, smoking duration, cigarettes smoked per day
LCRAT (2016)	LC Incidence	5 y	age, gender, race, education, BMI, pack-years, smoking duration, years since cessation, cigarettes smoked per day, history of COPD, family history of lung cancer (none, early, late onset)
LCDRAT (2016)	LC Mortality	5 y	age, gender, race, education, BMI, pack-years, smoking duration, years since cessation, cigarettes smoked per day, history of COPD, family history of lung cancer (none, early, late onset)

- Focused on Western populations (USA & Europe)
- Developed and validated using cohorts with >90% whites
- Post validation studies: limited to Western (White) populations
- Performance in Asian populations?

Asian Smokers

- Low smoking intensity
- Late-onset smoking
- Low rates of smoking cessation
- Very low prevalence of smoking in women
- High prevalence of never-smoking lung cancer

Study Aims

Asia: the major epicenter of lung cancer with more than 50% of lung cancers worldwide

- **Aim1:** To evaluate the statistical performance, *i.e.*, calibration and discrimination, of 11 lung cancer risk models in multiple Asia populations (using 19 prospective cohorts)
- **Aim2:** To better refine risk models for Asians by developing new prediction models incorporating Asian-specific risk estimates based on two well-characterized prospective cohorts (SMHS/SWHS)

The Asia Cohort Consortium

- 19 prospective cohorts in Asia
  - : 4 cohorts from China
  - : 7 cohorts from Japan
  - : 5 cohorts from South Korea
  - : 1 cohort from Taiwan
  - : 1 cohort from Mumbai
  - : 1 cohort from Iran
- Final analytic sample
  - : 186,458 Asian smokers
  - : ~6,800 incident lung cancer

https://www.asiacohort.org/

Methods

- **Study Populations**

  : 186,458 Asian ever-smokers (aged ≥50) from 19 ACC cohorts

- **Evaluation of 11 Western Lung Cancer Risk Prediction Models**

  : Calibration (E/O ratios) and discrimination (AUC)

  : Based on the two-stage random-effects meta-analysis method

  : Using the publicly available R package

- **Development of Shanghai Lung Cancer Risk Prediction Models**

  : Using data on ever-smokers aged 40-75 years at baseline within the SMHS and SWHS

  : Two cause-specific proportional hazards models, considering the competing mortality hazard

    - Shanghai lung cancer incidence (Shanghai-LCM) and death model (Shanghai-LCDM)

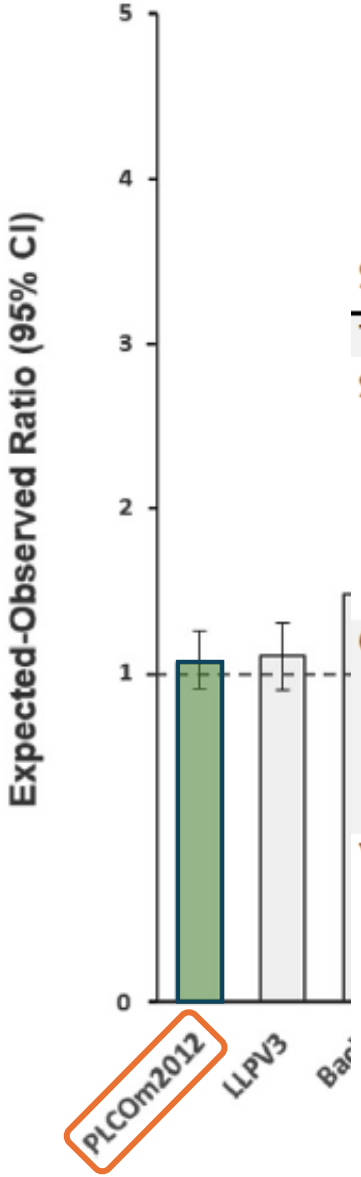
  : External validation using individual participant data from 17 ACC cohorts

Participating cohorts in the Asia Cohort Consortium

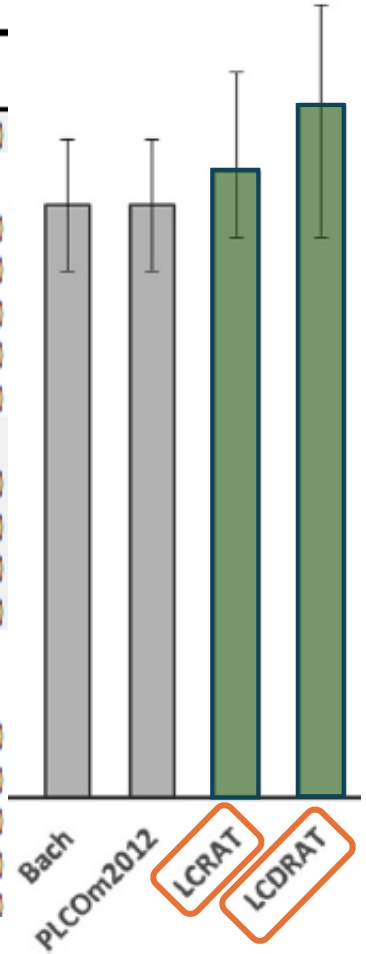
Participating Cohorts	No. of Participants[a]	Baseline Survey	Follow-up Years[b]	Age at Baseline[c]	Men (%)	Current Smokers (%)	Smoking Pack-Years[c]		Eligible USPSTF[d]		No. of Lung Cancer	
							Men	Women	2013 (%)	2021 (%)	Cases[e] (N)	Deaths (N)
Chinese												
SMHS	24,069	2002-2006	11.5	60.0	100.0	84.2	27.7	N.A.	27.1	59.4	845	695
SWHS	1584	1997-2000	15.4	63.6	0.0	81.6	N.A.	14.7	14.4	25.3	85	76
SCHS	15,816	1994-2005	11.5	64.4	82.7	53.7	31.4	15.0	35.1	52.7	906	791
SCS	8485	1986-1989	20.4	57.9	100.0	89.2	26.9	N.A.	32.0	61.0	823	801
CBCSP	3451	1991-1992	14.0	57.6	98.9	85.0	24.5	6.8	28.3	68.9	N.A.	92
Japanese												
JACC	22,699	1988-1990	14.5	62.1	90.8	64.2	28.7	14.4	34.7	54.9	1021	910
Miyagi	11,414	1990-1990	19.5	57.5	91.3	72.4	34.7	17.1	37.9	70.2	811	445
Ohsaki	16,026	1996-1996	10.3	64.2	90.0	63.9	34.8	15.4	45.6	67.3	722	486
3Pref Miyagi	6610	1984-1984	7.4	62.0	83.3	67.7	36.6	18.5	38.6	67.3	125	94
3Pref Aichi	10,374	1985-1985	11.1	61.8	81.7	64.0	38.3	18.4	37.5	66.0	317	283
LSS (RERF)	12,255	1963-1993	16.3	60.6	78.8	88.1	30.2	12.1	47.2	74.2	N.A.	574
Takayama	8369	1992-1992	12.5	63.3	83.3	58.4	25.1	11.8	23.1	40.4	302	N.A.
Korean												
KMCC	5218	1993-2004	12.1	63.5	82.0	68.7	33.4	13.7	41.7	63.7	307	257
Seoul	4818	1992-1993	15.4	54.0	100.0	62.3	24.6	N.A.	16.3	56.8	N.A.	65
KNCC	8278	2002-	8.9	57.3	94.0	39.2	24.4	9.6	20.5	48.8	155	43
Namwon	3356	2004-2007	11.6	64.2	90.2	85.5	31.7	15.8	42.0	64.2	172	127
KCS	3101	1985-1985	12.5	67.4	71.1	91.0	40.5	15.2	49.9	63.3	124	103
Indian												
Mumbai	16,093	1991-1997	4.8	60.0	99.0	73.7	12.9	5.9	7.0	15.8	52	52
Iranian												
GCS	4442	2003-2008	11.1	59.5	93.5	54.7	20.7	7.8	18.5	37.1	54	47
Total	186,458	1963-2008	12.7	61.1	89.7	68.8	28.3	14.5	31.8	55.8	6821	5941

Calibration & Discrimination of Western Models in Asian Populations

Extremely unstable among Asian
"low-intensity smokers" and
"long-term abstainers"



Stratification	Expected-to-Observed Ratio (95% CI)			AUC (95% CI)		
	$PLCO_{m2012}$	LCRAT	LCDRAT	$PLCO_{m2012}$	LCRAT	LCDRAT
Total study population	1.06 (0.90-1.25)	1.55 (1.30-1.86)	1.67 (1.39-2.00)	0.68 (0.66-0.70)	0.69 (0.67-0.72)	0.71 (0.67-0.74)
Smoking pack-years						
<10	0.19 (0.14-0.26)	1.26 (0.84-1.88)	1.38 (0.83-2.28)	0.58 (0.54-0.63)	0.65 (0.59-0.71)	0.72 (0.66-0.78)
10-19	0.62 (0.49-0.77)	1.33 (1.02-1.75)	1.45 (1.10-1.90)	0.62 (0.58-0.67)	0.65 (0.60-0.71)	0.69 (0.61-0.78)
20-29	1.05 (0.84-1.31)	1.16 (0.97-1.39)	1.05 (0.84-1.31)	0.67 (0.63-0.72)	0.69 (0.63-0.76)	0.71 (0.65-0.77)
30-39	1.12 (0.89-1.41)	1.49 (1.19-1.86)	1.57 (1.23-2.01)	0.64 (0.60-0.69)	0.65 (0.58-0.72)	0.64 (0.58-0.72)
≥40	1.34 (1.11-1.63)	1.67 (1.38-2.02)	1.68 (1.36-2.08)	0.61 (0.58-0.65)	0.63 (0.59-0.67)	0.62 (0.56-0.68)
Cigarettes smoked/d						
<10	0.24 (0.18-0.33)	1.59 (1.18-2.15)	1.90 (1.30-2.76)	0.63 (0.58-0.69)	0.72 (0.64-0.80)	0.81 (0.76-0.86)
10-19	1.00 (0.80-1.24)	1.35 (1.11-1.64)	1.39 (1.16-1.67)	0.70 (0.66-0.73)	0.69 (0.65-0.73)	0.68 (0.63-0.72)
20-29	1.22 (1.03-1.46)	1.46 (1.23-1.73)	1.55 (1.28-1.88)	0.68 (0.64-0.72)	0.69 (0.65-0.73)	0.71 (0.66-0.78)
≥30	1.25 (1.02-1.54)	1.56 (1.26-1.94)	1.61 (1.27-2.03)	0.68 (0.60-0.76)	0.69 (0.62-0.76)	0.71 (0.63-0.80)
Years since quitting smoking						
<5	0.98 (0.76-1.27)	1.50 (1.11-2.03)	1.26 (0.91-1.77)	0.70 (0.65-0.75)	0.67 (0.60-0.74)	0.76 (0.69-0.83)
5-9	0.99 (0.78-1.27)	1.08 (0.84-1.40)	0.85 (0.67-1.09)	0.71 (0.63-0.80)	0.67 (0.56-0.81)	0.76 (0.68-0.84)
10-14	0.83 (0.64-1.07)	0.89 (0.66-1.19)	0.93 (0.70-1.24)	0.67 (0.57-0.78)	0.69 (0.55-0.81)	0.62 (0.51-0.76)
15-19	0.63 (0.48-0.83)	0.73 (0.56-0.94)	0.77 (0.51-1.15)	0.67 (0.58-0.78)	0.67 (0.57-0.78)	0.64 (0.52-0.80)
≥20	0.48 (0.34-0.78)	0.55 (0.39-0.76)	0.60 (0.38-0.95)	0.64 (0.55-0.73)	0.71 (0.64-0.79)	0.68 (0.59-0.77)



Predictor	Definition	Model
age **agex** **agex2** **log_age** **lage2**	40-75 years at the baseline age-55 agex^2 log(age)-log(55) log_age^2	**Shanghai-LCM** Cox model lung cancer incidence cases: Surv(followed_years, lung_cancer_case) ~ log_age + gender + educatn1 + educatn2 + educatn4 + educatn5 + bmix + bmix2 + smkyrsc1 + smkyrsc2 + smkyrsc4 + smkyrsc5 + log_pkyrs + log_qtyrs + fam_lcnum. **Shanghai-LCDM** Cox model lung cancer deaths: Surv(followed_years, lung_cancer_death) ~ log_age + gender + educatn1 + educatn2 + educatn4 + educatn5 + bmix + bmix2 + smkyrsc1 + smkyrsc2 + smkyrsc4 + smkyrsc5 + log_pkyrs + log_qtyrs + fam_lcnum.
gender	female vs male 0-male (reference) 1-female	
educatn	education levels 1-elementary school or lower 2-middle school 3-high school (reference) 4-some college 5-college or higher	
bmix **bmix2** **log_bmi** **lbmi2**	bmi-25 bmix^2 log(bmi)-log(25) log_bmi^2	
copd	COPD history 0-no (reference) 1-yes	
smoke status	smoking status 1-former smoker (reference) 2-current smoker	
cigar per day **years_smoked** **smkyrsc**	number of cigarettes per day number of smoking years categorized number of smoking years 1-<10 years 2-<20 years 3-<30years (reference) 4-<40 years 5->= 40 years	
packyears **log_pkyrs**	cigar per day*years smoked/20 log(packyears+1)	
years_quitted **log_qtyrs**	duration (years) of smoking cessation log(years_quitted+1)	
fam lcnum (continuous)	number of lung cancer cases in family members 0-None 1-one lung cancer case in family members 2- ≥2 lung cancer cases in family members	

Development of Lung Cancer Risk Models Based on the SMHS & SWHS

To predict 1- to 10-year cumulative risk of developing lung cancer or dying from lung cancer, considering the competing mortality hazard

Stratification	Expected to Observed Ratio (95% CI)		AUC (95% CI)	
	Shanghai-LCM	Shanghai-LCDM	Shanghai-LCM	Shanghai-LCDM
Total study population	1.55 (1.24-1.93)	1.80 (1.44-2.25)	0.70 (0.67-0.72)	0.72 (0.69-0.74)
Ethnicity				
Chinese	0.98 (0.89-1.08)	1.08 (0.85-1.38)	0.70 (0.65-0.76)	0.69 (0.63-0.77)
Japanese	1.70 (1.36-2.13)	1.97 (1.52-2.55)	0.70 (0.66-0.75)	0.71 (0.67-0.75)
Korean	1.20 (0.73-1.99)	1.72 (1.11-2.66)	0.69 (0.66-0.72)	0.75 (0.68-0.81)
Indian	4.24 (3.00-6.00)	4.39 (2.94-6.55)	0.64 (0.55-0.74)	0.65 (0.54-0.76)
Iranian	2.75 (1.71-4.42)	2.31 (1.41-3.77)	0.74 (0.64-0.84)	0.75 (0.65-0.85)
Age, y				
50-59	1.67 (1.27-2.19)	1.79 (1.37-2.33)	0.68 (0.64-0.71)	0.67 (0.64-0.71)
60-69	1.45 (1.15-1.82)	1.64 (1.31-2.05)	0.65 (0.62-0.67)	0.66 (0.64-0.68)
≥70	1.65 (1.27-2.14)	1.89 (1.40-2.54)	0.71 (0.69-0.73)	0.71 (0.68-0.75)
Gender				
Men	1.52 (1.22-1.89)	1.77 (1.43-2.18)	0.69 (0.66-0.71)	0.70 (0.67-0.73)
Women	1.85 (1.21-2.83)	1.78 (1.06-2.98)	0.76 (0.69-0.84)	0.84 (0.78-0.91)
Smoking status				
Current	1.58 (1.25-1.99)	1.88 (1.47-2.40)	0.70 (0.68-0.71)	0.71 (0.68-0.74)
Former	1.40 (1.09-1.79)	1.45 (1.13-1.87)	0.69 (0.65-0.74)	0.70 (0.60-0.81)
Smoking pack-years				
<10	0.88 (0.64-1.22)	1.15 (0.78-1.70)	0.70 (0.63-0.78)	0.77 (0.71-0.83)
10-19	1.36 (1.00-1.86)	1.65 (1.21-2.26)	0.69 (0.64-0.74)	0.72 (0.65-0.80)
20-29	1.52 (1.19-1.93)	1.53 (1.17-2.01)	0.71 (0.65-0.78)	0.72 (0.65-0.79)
30-39	1.45 (1.10-1.91)	1.71 (1.27-2.30)	0.66 (0.63-0.70)	0.67 (0.62-0.71)
≥40	1.64 (1.30-2.06)	1.73 (1.35-2.21)	0.64 (0.60-0.67)	0.67 (0.62-0.71)
Cigarettes smoked/d				
<10	1.15 (0.79-1.67)	1.66 (1.03-2.67)	0.71 (0.63-0.80)	0.81 (0.75-0.87)
10-19	1.52 (1.19-1.94)	1.77 (1.38-2.26)	0.70 (0.66-0.74)	0.69 (0.65-0.74)
20-29	1.48 (1.20-1.82)	1.67 (1.33-2.11)	0.69 (0.66-0.72)	0.70 (0.64-0.77)
≥30	1.53 (1.18-1.99)	1.69 (1.29-2.22)	0.69 (0.62-0.78)	0.73 (0.66-0.82)
Years since quitting smoking				
<5	1.58 (1.12-2.24)	1.55 (1.02-2.35)	0.63 (0.53-0.74)	0.68 (0.58-0.80)
5-9	1.28 (0.95-1.73)	1.01 (0.78-1.32)	0.76 (0.67-0.87)	0.83 (0.76-0.91)
10-14	1.09 (0.83-1.42)	1.11 (0.81-1.52)	0.66 (0.54-0.80)	0.68 (0.55-0.84)
15-19	1.08 (0.79-1.48)	1.02 (0.66-1.59)	0.73 (0.68-0.79)	0.74 (0.64-0.86)
≥20	0.88 (0.65-1.19)	0.94 (0.67-1.32)	0.70 (0.63-0.77)	0.69 (0.63-0.76)

External Validation of Shanghai Models

Good internal validity

= Overall AUCs 0.78-0.80

Conclusion

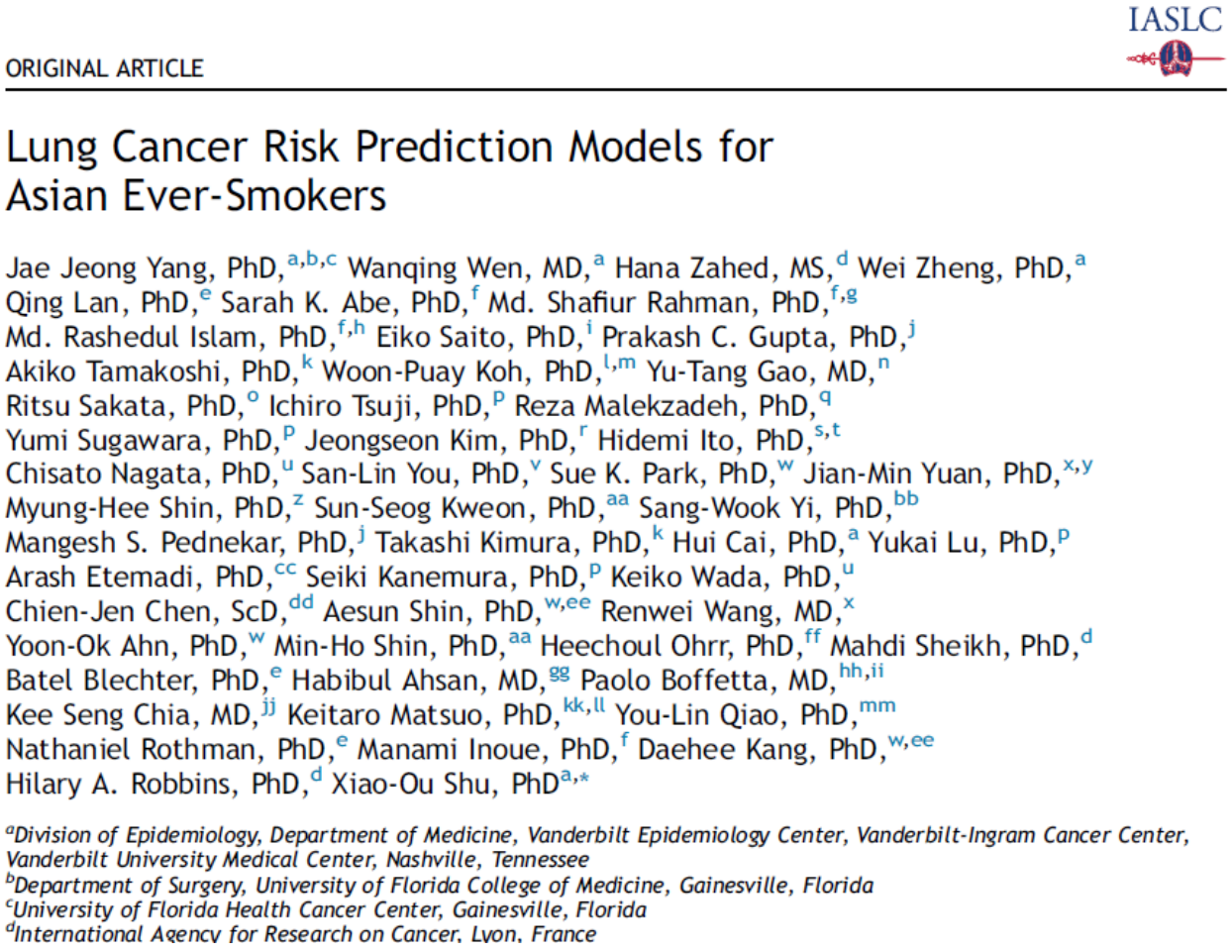
- **Lung cancer risk models developed in the U.S. and Europe**
  - $PLCO_{m2012}$, LCRAT, and LCDRAT had good predictive performance in Asian populations
  - Performed poorly in predicting lung cancer risk among Asians who reported low-intensity smoking or who had quit smoking for prolonged periods

- **Shanghai lung cancer risk prediction models**
  - *Improved predictive performance* for *low-intensity smokers* and *long-term quitters* who were particularly prevalent in Asia but not captured well by Western models
  - Also had room to be refined for universal application to diverse Asian populations

- Importance of incorporating *Asia-specific risk estimates into* ***personalized lung cancer risk assessment*** *to better implement risk-based LDCT screening in Asia*

- Further need for *country-specific adjustment* in identifying at-risk Asians who are most eligible for LDCT screening

Acknowledgement



IASLC

ORIGINAL ARTICLE

Lung Cancer Risk Prediction Models for Asian Ever-Smokers

Jae Jeong Yang, PhD,[a,b,c] Wanqing Wen, MD,[a] Hana Zahed, MS,[d] Wei Zheng, PhD,[a] Qing Lan, PhD,[e] Sarah K. Abe, PhD,[f] Md. Shafiur Rahman, PhD,[f,g] Md. Rashedul Islam, PhD,[f,h] Eiko Saito, PhD,[i] Prakash C. Gupta, PhD,[j] Akiko Tamakoshi, PhD,[k] Woon-Puay Koh, PhD,[l,m] Yu-Tang Gao, MD,[n] Ritsu Sakata, PhD,[o] Ichiro Tsuji, PhD,[p] Reza Malekzadeh, PhD,[q] Yumi Sugawara, PhD,[p] Jeongseon Kim, PhD,[r] Hidemi Ito, PhD,[s,t] Chisato Nagata, PhD,[u] San-Lin You, PhD,[v] Sue K. Park, PhD,[w] Jian-Min Yuan, PhD,[x,y] Myung-Hee Shin, PhD,[z] Sun-Seog Kweon, PhD,[aa] Sang-Wook Yi, PhD,[bb] Mangesh S. Pednekar, PhD,[j] Takashi Kimura, PhD,[k] Hui Cai, PhD,[a] Yukai Lu, PhD,[p] Arash Etemadi, PhD,[cc] Seiki Kanemura, PhD,[p] Keiko Wada, PhD,[u] Chien-Jen Chen, ScD,[dd] Aesun Shin, PhD,[w,ee] Renwei Wang, MD,[x] Yoon-Ok Ahn, PhD,[w] Min-Ho Shin, PhD,[aa] Heechoul Ohrr, PhD,[ff] Mahdi Sheikh, PhD,[d] Batel Blechter, PhD,[e] Habibul Ahsan, MD,[gg] Paolo Boffetta, MD,[hh,ii] Kee Seng Chia, MD,[jj] Keitaro Matsuo, PhD,[kk,ll] You-Lin Qiao, PhD,[mm] Nathaniel Rothman, PhD,[e] Manami Inoue, PhD,[f] Daehee Kang, PhD,[w,ee] Hilary A. Robbins, PhD,[d] Xiao-Ou Shu, PhD[a,*]

[a]*Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University Medical Center, Nashville, Tennessee*
[b]*Department of Surgery, University of Florida College of Medicine, Gainesville, Florida*
[c]*University of Florida Health Cancer Center, Gainesville, Florida*
[d]*International Agency for Research on Cancer, Lyon, France*

J Thorac Oncol. 2024 Mar;19(3):451-464.

VANDERBILT UNIVERSITY MEDICAL CENTER

Dr. Xiao-Ou Shu

Dr. Wanqing Wen

Dr. Wei Zheng

Dr. Hui Cai

International Agency for Research on Cancer

World Health Organization

Dr. Hilary Robbins

Dr. Hana Zahed

The Asia Cohort Consortium

All contributing cohorts

ACC Coordinating Center

THANK YOU

yang.jae@surgery.ufl.edu