Evaluation of risk prediction models to select lung cancer screening participants in Europe: a prospective cohort consortium analysis

Xiaoshuang Feng, Patrick Goodley, Karine Alcala, Florence Guida, Rudolf Kaaks, Roel Vermeulen, George S Downward, Catalina Bonet, Sandra M Colorado-Yohar, Demetrius Albanes, Stephanie J Weinstein, Marcel Goldberg, Marie Zins, Caroline Relton, Arnulf Langhammer, Anne Heidi Skogholt, Mattias Johansson, Hilary A Robbins

Summary

Background Lung cancer risk prediction models might efficiently identify individuals who should be offered lung cancer screening. However, their performance has not been comprehensively evaluated in Europe. We aimed to externally validate and evaluate the performance of several risk prediction models that predict lung cancer incidence or mortality in prospective European cohorts.

Methods We analysed 240 137 participants aged 45–80 years with a current or former smoking history from nine European countries in four prospective cohorts from the pooled database of the Lung Cancer Cohort Consortium: the Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study (Finland), the Nord-Trøndelag Health Study (Norway), CONSTANCES (France), and the European Prospective Investigation into Cancer and Nutrition (Denmark, Germany, Italy, Spain, Sweden, the Netherlands, and Norway). We evaluated ten lung cancer risk models, which comprised the Bach, the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial 2012 model ($PLCO_{m2012}$), the Lung Cancer Risk Assessment Tool (LCRAT), the Lung Cancer Death Risk Assessment Tool (LCDRAT), the Nord-Trøndelag Health Study (HUNT), the Optimized Early Warning Model for Lung Cancer Risk (OWL), the University College London—Death (UCLD), the University College London—Incidence (UCLI), the Liverpool Lung Project version 2 (LLP version 2), and the Liverpool Lung Project version 3 (LLP version 3) models. We quantified model calibration as the ratio of expected to observed cases or deaths and discrimination using the area under the receiver operating characteristic curve (AUC). For each model, we also identified risk thresholds that would screen the same number of individuals as each of the US Preventive Services Task Force 2021 (USPSTF-2021), the US Preventive Services Task Force 2013 (USPSTF-2013), and the Nederlands–Leuvens Longkanker Screenings Onderzoek (NELSON) criteria.

Findings Among the participants, 1734 lung cancer cases and 1072 lung cancer deaths occurred within five years of enrolment. Most models had reasonable calibration in most countries, although the LLP version 2 overpredicted risk by more than 50% in eight countries (expected to observed ≥1·50). The $PLCO_{m2012}$, LCDRAT, LCRAT, Bach, HUNT, OWL, UCLD, and UCLI models showed similar discrimination in most countries, with AUCs ranging from 0·68 (95% CI 0·59–0·77) to 0·83 (0·78–0·89), whereas the LLP version 2 and LLP version 3 showed lower discrimination, with AUCs ranging from 0·64 (95% CI 0·57–0·72) to 0·78 (0·74–0·83). When pooling data from all countries (but excluding the HUNT cohort), 33·9% (73 313 of 216 387) of individuals were eligible by USPSTF-2021 criteria, which included 74·8% (1185) of lung cancers and 76·3% (730) of lung cancer deaths occurring over 5 years. Fewer individuals were selected by USPSTF-2013 and NELSON criteria. After applying thresholds to select a population of equal size to USPSTF-2021, the $PLCO_{m2012}$, LCDRAT, LCRAT, Bach, HUNT, OWL, UCLD, and UCLI, models identified 77·6%–79·1% of future cases, although they selected slightly older individuals compared with USPSTF-2021 criteria. Results were similar for USPSTF-2013 and NELSON.

Interpretation Several lung cancer risk prediction models showed good performance in European countries and might improve the efficiency of lung cancer screening if used in place of categorical eligibility criteria.

Funding US National Cancer Institute, l'Institut National du Cancer, Cancer Research UK.

Lancet Digit Health 2024; 6: e614–24

Genomic Epidemiology Branch (X Feng PhD, K Alcala MS, M Johansson PhD, H A Robbins PhD), **Environment and Lifestyle Epidemiology Branch** (F Guida PhD), **International Agency for Research on Cancer, Lyon, France; Division of Immunology, Immunity to Infection and Respiratory Medicine, University of Manchester, Manchester, UK** (P Goodley MBBCh); **Manchester Thoracic Oncology Centre, Manchester University NHS Foundation Trust, Manchester, UK** (P Goodley); **Department of Cancer Epidemiology, German Cancer Research Center, Heidelberg, Germany** (Prof R Kaaks PhD); **Translational Lung Research Center Heidelberg, Member of the German Center for Lung Research (DZL), Heidelberg, Germany** (Prof R Kaaks); **Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, Netherlands** (Prof R Vermeulen PhD, G S Downward PhD); **Department of Population Health Sciences, Institute for Risk Assessment Sciences, University of Utrecht, Utrecht, Netherlands** (Prof R Vermeulen, G S Downward); **Nutrition and Cancer Group, Epidemiology, Public Health, Cancer Prevention and Palliative Care Program, Bellvitge Biomedical Research Institute, L'Hospitalet de Llobregat, Barcelona, Spain** (C Bonet MSc); **Unit of Nutrition and Cancer, Catalan Institute of Oncology, L'Hospitalet de Llobregat, Barecelona, Spain** (C Bonet); **Department of Epidemiology, Murcia Regional Health Council, IMIB-Arrixaca,**

Murcia, Spain (S M Colorado-Yohar PhD); **CIBER Epidemiología y Salud Pública, Madrid, Spain** (S M Colorado-Yohar); **Research Group on Demography and Health, National Faculty of Public Health, University of Antioquia, Medellín, Colombia** (S M Colorado-Yohar); **Metabolic Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD, USA** (D Albanes MD, S J Weinstein PhD); **Population-based Epidemiological Cohorts Unit, INSERM UMS 11, Villejuif, France** (Prof M Goldberg PhD, Prof M Zins PhD); **Paris Cité University, Paris, France** (Prof M Goldberg, Prof M Zins); **MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK** (Prof C Relton PhD); **School of Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK** (Prof C Relton); **HUNT Research Center, Department of Public Health and Nursing, Norwegian University of Science and Technology, Levanger, Norway** (Prof A Langhammer PhD); **Levanger Hospital, Nord-Trøndelag Hospital Trust, Levanger, Norway** (Prof A Langhammer); **Department of Public Health and Nursing, KG Jebsen Centre for Genetic Epidemiology, Norwegian University of Science and Technology, Trondheim, Norway** (A H Skogholt PhD)

Correspondence to: Dr Hilary A Robbins, Genomic Epidemiology Branch, International Agency for Research on Cancer, 69366 Lyon CEDEX 07, France
robbinsh@iarc.who.int

Research in context

Evidence before this study

Implementation of lung cancer screening is rapidly proceeding in Europe, and risk prediction models might identify candidates for screening more efficiently than standard age and smoking criteria. We did a PubMed search for articles published in English up to March 1, 2023, including the terms "lung cancer" and "risk prediction" and ("performance" or "validation"). Among 191 papers retrieved, we confirmed that previous studies have not comprehensively evaluated whether existing lung cancer risk prediction models have sufficient performance to be used in European populations, nor identified which model(s) identify individuals at high risk most efficiently in Europe.

Added value of this study

This study used the large, pooled database of the Lung Cancer Cohort Consortium to externally validate and directly compare the performance of ten risk prediction models (Bach, $PLCO_{m2012}$, LCRAT, LCDRAT, HUNT, OWL, UCLD, UCLI, LLP version 2, and LLP version 3) that might be considered for use in defining eligibility for lung cancer screening in nine European countries. With over 240 000 current or former smoking participants included, this is the first study to comprehensively compare the currently available lung cancer risk prediction models across Europe. The results provide further evidence that risk prediction models outperform categorical criteria for lung cancer screening eligibility, and highlight that multiple models are probably appropriate for use in the European context.

Implications of all the available evidence

For lung cancer screening eligibility in Europe, risk prediction models can identify more future lung cancer cases and deaths than categorical eligibility criteria, without screening more people. A group of risk models developed by use of data from the USA and Europe have similar overall performance, and any one of them is probably reasonable to use. For the ongoing implementation of lung cancer screening in Europe, it is advantageous to pilot the use of a risk prediction model to define who is eligible for screening, with the choice among the well-performing models depending on specific considerations for a given context.

Introduction

Screening by low-dose CT can reduce lung cancer mortality among people with a heavy smoking history.[1-4] Pilot programmes are now proceeding in Europe,[5] and the European Commission has recommended implementation of lung cancer screening with a stepwise approach.[6] However, questions remain about how to optimally target screening to those individuals who are most likely to benefit. Several risk prediction models have been developed to predict an individual's risk of lung cancer or lung cancer mortality. Compared with eligibility criteria that use categories of age and smoking information, these models might identify future lung cancer cases and deaths with greater efficiency.[7]

However, use of a poorly performing risk model to define screening eligibility could lead to an inefficient screening programme with a poor balance of benefits and harms. The Bach model, Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial 2012 model ($PLCO_{m2012}$), Lung Cancer Risk Assessment Tool (LCRAT), and Lung Cancer Death Risk Assessment Tool (LCDRAT) have been shown to perform well in the US population.[7] Considering the diverse demographic and smoking profiles across Europe,[8] it is possible that established models perform differently in Europe. These risk models have been only partially tested in Europe (the UK, Germany, and Finland),[9–11] and information for other countries is lacking.

There are also various newly developed models. In particular, the Norwegian Nord-Trøndelag Health Study (HUNT) model has only been validated in one small cohort,[12] but is being considered for use in the Nordic countries. The Optimized Early Warning Model for Lung Cancer Risk (OWL), University College London—Incidence (UCLI), and University College London—Death (UCLD) models are newly developed.[13,14] The Liverpool Lung Project (LLP) version 2 model is used in the UK but a new version, LLP version 3, has been proposed.[15] A thorough evaluation of the performance of available lung cancer risk models in European countries is required before they can be implemented within lung cancer screening programmes.

The Lung Cancer Cohort Consortium (LC3) database includes detailed risk factor and lung cancer outcomes data from 24 prospective cohort studies in North America, Asia, Australia, and Europe,[16] and was used to evaluate lung cancer risk models in the UK.[9] In the current study, we analysed data from nine European countries in LC3 to evaluate the performance of ten models that predict lung cancer incidence or mortality.

Methods

Study design and participants

We analysed data from four prospective cohorts in which participants were enrolled and followed up prospectively for cancer incidence and outcomes: the Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study (ATBC-Finland), HUNT (Norway), CONSTANCES (France), and the European Prospective Investigation into Cancer and Nutrition (EPIC, ten countries).[17] The ATBC study was a randomised trial including 29133 Finnish male participants with a smoking history aged 49–70 years at recruitment (1985–88). The HUNT study is a population-based cohort including 125000 Norwegian participants aged 20 years or older from three subcohorts: HUNT1

(1984–86), HUNT2 (1995–97) and HUNT3 (2006–08). We included participants from HUNT2 and HUNT3 (n=76 006). The EPIC study is a prospective cohort of 521 330 participants mostly between 35 years and 70 years of age enrolled at 23 centres in ten western European countries (Denmark, France, Germany, Greece, Italy, Spain, Sweden, the Netherlands, Norway, and the UK) between 1992 and 2000. The CONSTANCES study is a population-based cohort of 202 045 French participants aged 18–69 years recruited from 2012 to 2020.[18,19] Data on demographics, smoking, other risk factors, and lung cancer outcomes were harmonised according to the LC3 protocol.[16] Lung cancer risk factor data in the four cohorts were primarily collected with questionnaire surveys. Lung cancer incidence data from the four cohorts were mainly collected from cancer registries in each country. Mortality outcomes including lung cancer mortality data from the four cohorts were collected through mortality registries in each country or active follow-up and death-record collection.

From all participants in these cohorts, we restricted to those who were aged 45–80 years and currently or formerly smoked at enrolment, including 29 133 in ATBC (Finland), 48 908 in CONSTANCES (France), 37 002 in EPIC-Denmark, 18 787 in EPIC-Germany, 17 483 in EPIC-Italy, 9499 in EPIC-Spain, 23 425 in EPIC-Sweden, 16 583 in EPIC-Netherlands, 15 567 in EPIC-Norway, and 23 750 in HUNT (Norway; 240 137 in total). We did not include EPIC-UK because we have previously published results for this cohort,[9] EPIC-France because the sample does not represent the general population,[17] or EPIC-Greece because of administrative and data use restrictions.

Statistical analysis

To evaluate the predictive performance of lung cancer risk models, we applied the previously developed algorithms for each model to calculate risk predictions for each individual. These predictions are the estimated probability (0–100%) of being diagnosed with lung cancer, or dying from lung cancer, over a specified time period. No modifications to the original models are made in this process. We evaluated ten lung cancer risk prediction models: the Bach model,[20] $PLCO_{m2012}$,[21] LCRAT, LCDRAT,[22] LLP version 2, LLP version 3,[15] HUNT,[23] OWL,[14] UCLD, and UCLI models.[13] Each model predicts risk of incident lung cancer, except LCDRAT and UCLD, which predict lung cancer mortality risk. The predictors include demographic information, smoking (smoking status, duration, cigarettes per day, pack-years, quit-years, etc), and health-related factors (details in appendix 1 p 1). Four models (Bach, $PLCO_{m2012}$, LCRAT, and LCDRAT) were developed with data from the USA, four models (LLP version 2, LLP version 3, HUNT, and OWL) with data from Europe including the UK, and UCLD and UCLI with data from both the USA and the UK. LLP version 2 and LLP version 3 were developed with case–control data combined with population lung cancer rates, and the remaining eight models that used prospective cohorts were developed (including randomised controlled trials) with 18 000–323 000 participants. In terms of statistical methodology, Bach, LCRAT, and LCDRAT are Cox proportional hazards models, $PLCO_{m2012}$, LLP version 2, and LLP version 3 are logistic regression models, and HUNT used both methods (here we evaluate the logistic regression model). UCLD and UCLI are ensemble machine learning models, and OWL used a single machine learning method (XGBoost).

We used multiple imputation for missing data. We applied predictive mean matching with chained equations to generate 15 imputed datasets, stratifying the imputation by cohort and smoking status as previously described.[16] In brief, we imputed missing values of education, BMI, personal history of cancer, chronic obstructive pulmonary disease (COPD), diabetes, family history of lung cancer, number of cigarettes smoked per day, ages at smoking initiation and cessation (table 1; appendix 1 p 2), daily cough, and secondhand smoke hours. Data on participant race and ethnicity are generally not collected in European cohorts. For risk models requiring race and ethnicity information, we coded participants in EPIC-Spain as Hispanic and all others as non-Hispanic White. For variables missing less than 35% within a group defined by cohort and smoking status (current or former), we imputed them using data from the same cohort. For those missing more than 35% within a group defined by cohort and smoking status, including those missing 100%, we imputed them using different cohorts that had low amounts of missing data, from the same geographical region whenever possible. When cause of death was missing, we assumed that people who died with a previous lung cancer diagnosis had died from lung cancer.

For $PLCO_{m2012}$ (6-year prediction time horizon), LCRAT (5-year) and LCDRAT (5-year), estimates of absolute lung cancer risk were calculated by use of the lcmodels package in R, which was created to validate lung cancer prediction models developed before 2018.[7,24] For the Bach model, we used code adapted from the lcmodels package to reduce the time horizon from 10 years to 5 years. For the LLP version 2 (5-year horizon), LLP version 3 (5-year)[15] and HUNT (6-year) models,[23] we developed new R scripts based on published papers. For OWL,[14] we applied the publicly available script in R and extracted the 5-year prediction for the current analysis. For UCLD (5-year) and UCLI (5-year), we applied the publicly available script in Python.[13]

We quantified calibration, which is a model's ability to predict the correct number of incident events, as the ratio of expected to observed lung cancer cases or deaths. The expected number of events was calculated by summing each model's predicted risks across cohort participants. We quantified discrimination, which is a model's ability to predict higher risk for events (lung cancer cases or deaths) compared with non-events,

See Online for appendix 1

using the area under the receiver operating characteristic curve (AUC) statistic. Rubin's rule was used to produce pooled estimates of expected to observed and AUC values from the 15 imputed datasets.[25] We excluded the HUNT cohort from validation of the HUNT model, since it would not provide external validation. We did stratified analyses for calibration and discrimination by sex (men and women), smoking (former and current), age (<55 years and ≥ 55 years), and education level (less than high school and high school and above).

We compared the performance of the ten risk models for selecting individuals for lung cancer screening versus the US Preventive Services Task Force 2021 criteria (USPSTF-2021: age 50–80 years, current or former smokers quit ≤15 years, ≥20 pack-years),[26] USPSTF-2013 criteria (age 55–80 years, current or former smokers quit ≤15 years, ≥30 pack-years),[27] and Nederlands–Leuvens Longkanker Screenings Onderzoek (NELSON) criteria (age 50–74, current or former smokers quit ≤10 years, >15 cigarettes per day for >25 years or >10 cigarettes per day for >30 years).[4]

	ATBC-Finland	CONSTANCES-France	EPIC-Denmark	EPIC-Germany	EPIC-Italy	EPIC-Spain	EPIC-Sweden	EPIC-Netherlands	EPIC-Norway	HUNT-Norway
All participants	n=29 133	n=48 908	n=37 002	n=18 787	n=17 483	n=9499	n=23 425	n=16 583	n=15 567	n=23 750
Lung cancer cases										
5-year	654 (2·2%)	231 (0·5%)	254 (0·7%)	106 (0·6%)	68 (0·4%)	40 (0·4%)	120 (0·5%)	64 (0·4%)	48 (0·3%)	149 (0·6%)
6-year	830 (2·8%)	256 (0·5%)	315 (0·9%)	143 (0·8%)	82 (0·5%)	41 (0·4%)	151 (0·6%)	84 (0·5%)	64 (0·4%)	190 (0·8%)
Lung cancer deaths										
5-year	414 (1·4%)	69 (0·1%)	188 (0·5%)	67 (0·4%)	49 (0·3%)	23 (0·2%)	93 (0·4%)	31 (0·2%)	23 (0·1%)	115 (0·5%)
USPSTF-2021										
Eligible	24 347 (83·6%)	7646 (15·6%)	14 467 (39·1%)	4012 (21·4%)	3590 (20·5%)	2809 (29·6%)	0	4465 (26·9%)	1611 (10·3%)	5230 (22·0%)
Ineligible	4786 (16·4%)	37 075 (75·8%)	16 720 (45·2%)	13 515 (71·9%)	12 551 (71·8%)	6542 (68·9%)	5029 (21·5%)	11 050 (66·6%)	13 428 (86·3%)	15 001 (63·2%)
Missing	0	4187 (8·6%)	5815 (15·7%)	1260 (6·7%)	1342 (7·7%)	148 (1·6%)	18 396 (78·5%)	1068 (6·4%)	528 (3·4%)	3519 (14·8%)
USPSTF-2013										
Eligible	12 794 (43·9%)	3465 (7·1%)	4705 (12·7%)	1628 (8·7%)	1119 (6·4%)	1188 (12·5%)	0	1486 (9·0%)	80 (0·5%)	1853 (7·8%)
Ineligible	16 339 (56·1%)	42 208 (86·3%)	28 708 (77·6%)	16 370 (87·1%)	15 671 (89·6%)	8218 (86·5%)	11 122 (47·5%)	14 470 (87·3%)	15 436 (99·2%)	18 857 (79·4%)
Missing	0	3235 (6·6%)	3589 (9·7%)	789 (4·2%)	693 (4·0%)	93 (1·0%)	12 303 (52·5%)	627 (3·8%)	51 (0·3%)	3040 (12·8%)
NELSON										
Eligible	21 476 (73·7%)	5705 (11·7%)	13 929 (37·6%)	3616 (19·2%)	3223 (18·4%)	2422 (25·5%)	0	4099 (24·7%)	2174 (14·0%)	3420 (14·4%)
Ineligible	7657 (26·3%)	37 900 (77·5%)	17 258 (46·6%)	13 911 (74·0%)	12 918 (73·9%)	6929 (72·9%)	5029 (21·5%)	11 416 (68·8%)	12 865 (82·6%)	17 388 (73·2%)
Missing	0	5303 (10·8%)	5815 (15·7%)	1260 (6·7%)	1342 (7·7%)	148 (1·6%)	18 396 (78·5%)	1068 (6·4%)	528 (3·4%)	2942 (12·4%)
Age, years										
45–49	47 (0·2%)	9888 (20·2%)	0	4796 (25·5%)	5430 (31·1%)	3749 (39·5%)	5029 (21·5%)	3184 (19·2%)	7656 (49·2%)	5126 (21·6%)
50–59	19 662 (67·5%)	19 297 (39·5%)	26 411 (71·4%)	9251 (49·2%)	9296 (53·2%)	4216 (44·4%)	10 380 (44·3%)	9852 (59·4%)	7911 (50·8%)	7916 (33·3%)
60–69	9391 (32·2%)	18 773 (38·4%)	10 591 (28·6%)	4740 (25·2%)	2718 (15·5%)	1534 (16·1%)	6620 (28·3%)	3545 (21·4%)	0	6290 (26·5%)
70–80	33 (0·1%)	950 (1·9%)	0	0	39 (0·2%)	0	1396 (5·96%)	2 (0·012%)	0	4418 (18·6%)
Median (IQR)	57 (53–61)	58 (51–64)	56 (52–60)	54 (49–60)	53 (48–57)	51 (47–57)	55 (50–61)	54 (50–59)	50 (47–52)	58 (51–68)
Sex										
Men	29 133 (100%)	26 787 (54·8%)	20 155 (54·5%)	11 685 (62·2%)	7552 (43·2%)	7190 (75·7%)	11 247 (48·0%)	3632 (21·9%)	0	13 469 (56·7%)
Women	0	22 121 (45·2%)	16 847 (45·5%)	7102 (37·8%)	9931 (56·8%)	2309 (24·3%)	12 178 (52·0%)	12 951 (78·1%)	15 567 (100%)	10 281 (43·3%)
Smoking status										
Current	29 133 (100%)	12 357 (25·3%)	17 326 (46·8%)	6466 (34·4%)	8312 (47·5%)	4740 (49·9%)	9892 (42·2%)	6929 (41·8%)	7937 (51·0%)	10 743 (45·2%)
Former, <15 quit years	0	12 206 (25·0%)	6979 (18·9%)	4715 (25·1%)	5478 (31·3%)	3176 (33·4%)	5766 (24·6%)	4220 (25·4%)	3463 (22·2%)	4482 (18·9%)
Former, ≥15 quit years	0	23193 (47·4%)	7234 (19·6%)	7288 (38·8%)	3659 (20·9%)	1523 (16·0%)	6563 (28·0%)	5209 (31·4%)	3836 (24·6%)	7047 (29·7%)
Missing	0	1152 (2·4%)	5463 (14·8%)	318 (1·7%)	34 (0·2%)	60 (0·6%)	1204 (5·1%)	225 (1·4%)	331 (2·1%)	1478 (6·2%)

Data are n (%). Data provided in this table are before imputation of missing data. For results in the following tables and figures, missing data were imputed as described in the Methods. ATBC=Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study. EPIC=European Prospective Investigation into Cancer and Nutrition. HUNT=Nord-Trøndelag Health Study. NELSON=Nederlands-Leuvens Longkanker Screenings Onderzoek. USPSTF-2013=US Preventive Services Task Force 2013. USPSTF-2021=US Preventive Services Task Force 2021.

Table 1: **Characteristics of Lung Cancer Cohort Consortium participants in nine European countries who were aged 45–80 years and currently or formerly smoked at enrolment**

For this analysis we used a single, randomly selected imputed dataset from among the 15 imputed datasets and again excluded the HUNT cohort. To appropriately compare the sensitivity of different eligibility criteria to identify future cases, it was necessary to specify that they would screen the same number of people. Therefore, for every comparison between each model and each of the three categorical approaches, we identified a threshold for the model's predicted risk values that would select an equally sized screening population as the categorical approach, whereby individuals with predicted risk above the threshold were considered screening-eligible. Next, we compared the number of lung cancer cases and deaths over 5 years that would be classified as screening-eligible by each model, where a higher number of cases or deaths implies more efficient screening.

In addition, we used Venn diagrams to present the number of participants and lung cancer cases selected by the LCDRAT, HUNT, and $PLCO_{m2012}$ risk models compared with categorical criteria (USPSTF or NELSON). We also presented the predicted risk, socioeconomics, and smoking characteristics of screening-eligible populations who were identified by categorical criteria overall, and who were exclusively identified by $PLCO_{m2012}$, LCDRAT, and HUNT models but not categorical criteria.

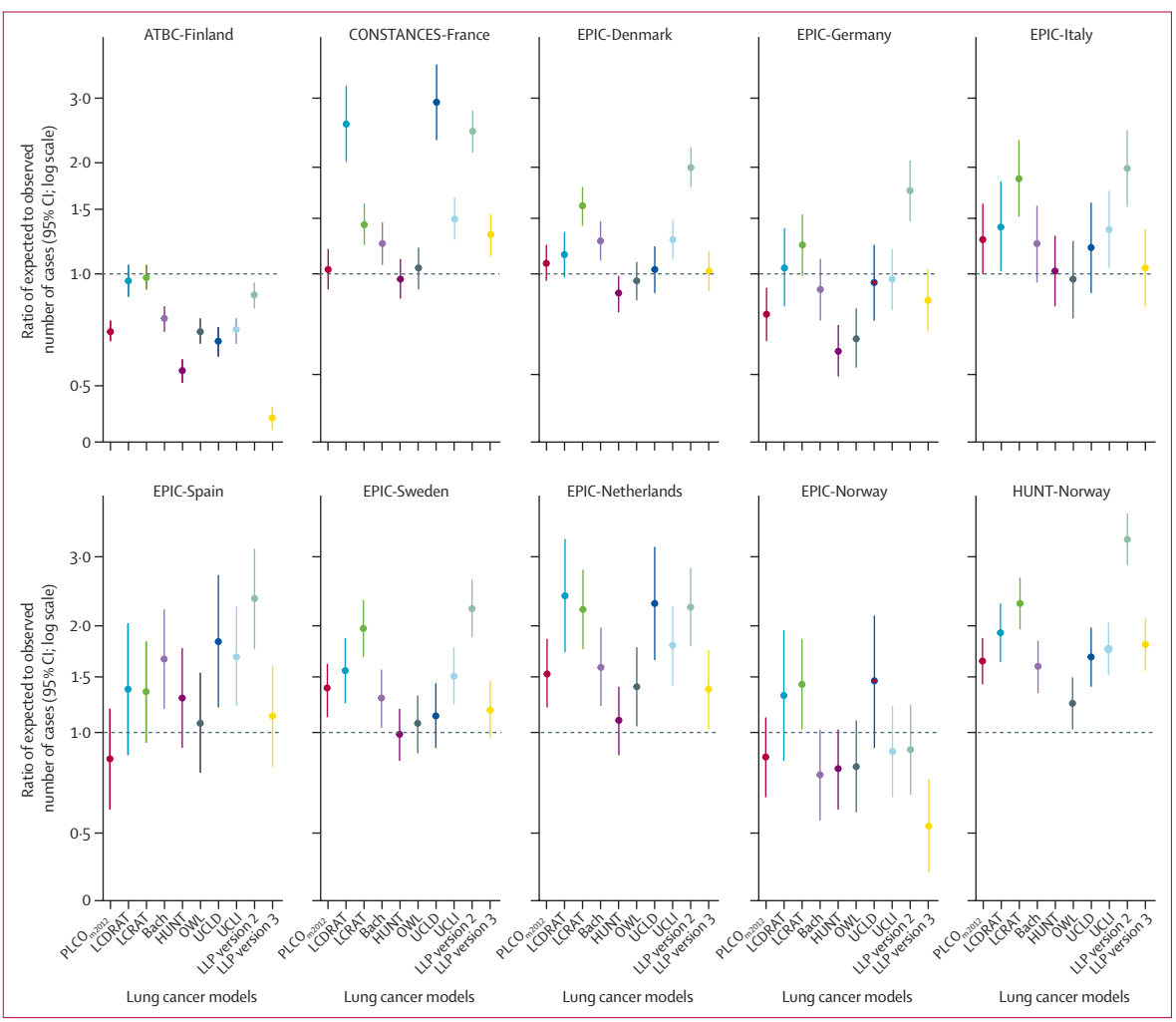

Figure 1: **Calibration of ten lung cancer risk prediction models in nine European countries, as measured by the ratio of expected to observed lung cancer cases or deaths**

Error bars are 95% CI. The optimal expected to observed value is 1·0 (perfect calibration; dashed line). 15 imputations were used for missing data and Rubin's rule was used to produce pooled expected to observed estimates from the 15 imputed datasets. The time horizons for each model are as follows: 5 years: LCDRAT, LCRAT, Bach, LLP version 2, LLP version 3, OWL, UCLD, UCLI; 6 years: $PLCO_{m2012}$ and HUNT models. ATBC=Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study. EPIC=European Prospective Investigation into Cancer and Nutrition. HUNT=Nord-Trøndelag Health Study. LCDRAT=Lung Cancer Death Risk Assessment Tool. LCRAT=Lung Cancer Risk Assessment Tool. LLP version 2=Liverpool Lung Project version 2. LLP version 3=Liverpool Lung Project version 3. OWL=Optimized Early Warning Model for Lung Cancer Risk. $PLCO_{m2012}$=Prostate, Lung, Colorectal, and Ovarian Cancer ScreeningTrial 2012 model. UCLD=University College London—Death. UCLI=University College London—Incidence.

Role of the funding source
The funders had no role in study design, data analysis, data interpretation, writing of the report, or the decision to submit.

Results

Among the 240 137 participants with a current or former smoking history aged 45–80 years in nine European countries, 1734 lung cancer cases and 1072 lung cancer deaths occurred within 5 years of enrolment (table 1). Median follow-up time ranged from 4·5 years (CONSTANCES in which 57% of participants were followed up for <5 years) to 18 years (ATBC). The representation of individuals aged 70–80 years was very small in most cohorts. The 5-year cumulative lung cancer incidence was highest in ATBC (2·2%), a cohort of currently smoking males, and lowest in EPIC-Norway (0·3%), a cohort of females younger than 60 years. Distributions of demographic variables, cigarettes per day, and smoking duration differed across countries (appendix 1 p 2).

Most models had reasonable calibration in most countries with most expected to observed ratios ranging from 0·70 to 1·50 (figure 1). The LCDRAT, LCRAT, and LLP version 2 models were more likely to overpredict risk than other models, with LLP version 2 overpredicting risk by more than 50% (expected to observed ≥1·5) in eight countries. Conversely, the HUNT model under-predicted risk by more than 20% (expected to observed ≤0·80) in

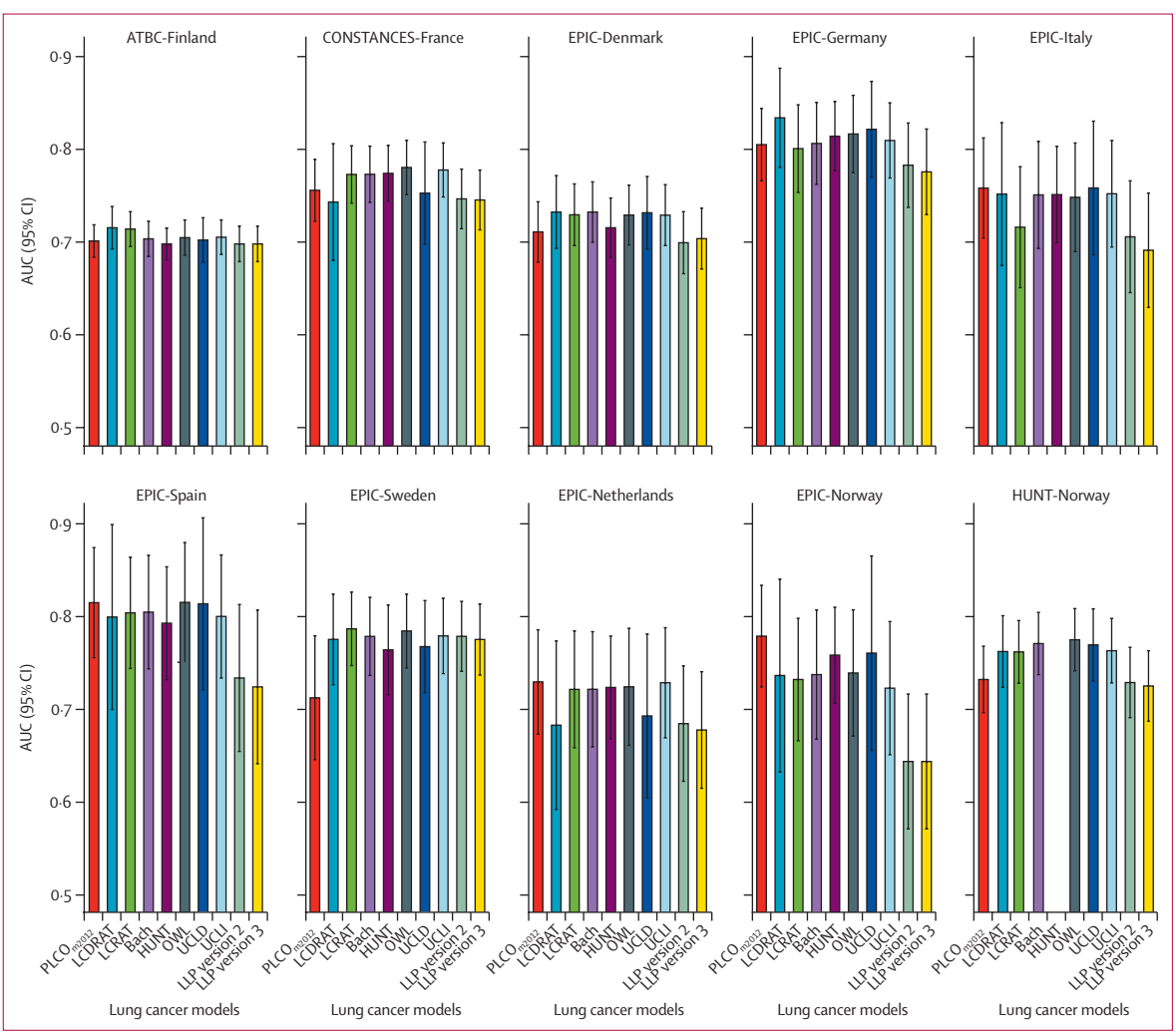

Figure 2: **Discrimination of ten lung cancer risk prediction models in nine European countries, as measured by the AUC**
Error bars are 95% CI. Higher AUC values indicate better risk discrimination (maximum value 1·0). AUCs are affected by the amount of variance in the model predictors, which differed substantially across countries and cohorts. 15 imputations were used for missing data and Rubin's rule was used to produce pooled AUC estimates from the 15 imputed datasets. The time horizons for each model are as follows: 5 years for the LCDRAT, LCRAT, Bach, LLP version 2, LLP version 3, OWL, UCLD, and UCLI models; 6 years for the $PLCO_{m2012}$ and HUNT models. AUC=area under the receiver operating characteristic curve. ATBC=Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study. EPIC=European Prospective Investigation into Cancer and Nutrition. HUNT=Nord-Trøndelag Health Study. LCDRAT=Lung Cancer Death Risk Assessment Tool. LCRAT=Lung Cancer Risk Assessment Tool. LLP version 2=Liverpool Lung Project version 2. LLP version 3=Liverpool Lung Project version 3. OWL=Optimized Early Warning Model for Lung Cancer Risk. $PLCO_{m2012}$=Prostate, Lung, Colorectal, and Ovarian 2012 model. UCLD=University College London—Death. UCLI=University College London—Incidence.

three countries. All models underpredicted risk in ATBC-Finland (expected to observed ranging from 0·41 [95% CI 0·38–0·44] to 0·98 [0·91–1·06]) and overpredicted in HUNT-Norway (expected to observed ranging from 1·20 [1·02–1·41] to 3·32 [2·83–3·90]) and in EPIC-Netherlands (expected to observed ranging from 1·08 [0·87–1·33] to 2·34 [1·65–3·33]). The two mortality prediction models, UCLD and LCDRAT, appeared to overestimate risk substantially in CONSTANCES-France whereas most incidence models did not. Stratified analyses for calibration by sex, smoking, age, and education are shown in appendix 2 p 1.

Model discrimination varied across countries (figure 2). This was expected, because AUCs are affected by the variance in model predictors, and the variance in age and smoking differed across the countries and cohorts in our study. Across the models in each country, the AUCs for LLP version 2 and LLP version 3 were typically lower compared with other models, with AUCs ranging from 0·64 (95% CI 0·57–0·72) to 0·78 (0·74–0·83), whereas the differences across the remaining eight models were modest, with AUCs ranging from 0·68 (95% CI 0·59–0·77) to 0·83 (0·78–0·89). Across the countries, AUCs in EPIC-Germany were highest (AUCs ranging from 0·78 to 0·83), followed by EPIC-Spain (AUCs ranging from 0·72 to 0·81), whereas the AUCs in ATBC-Finland were lower than the other countries (AUCs ranging from 0·70 to 0·72). Stratified analyses for discrimination by sex, smoking, age, and education are shown in appendix 1 (pp 4–5).

In the combined population of 216 387 participants after excluding the HUNT cohort, 73 313 (33·9%) individuals were eligible for screening by USPSTF-2021 criteria, a group that included 1185 (74·8%) of 1585 lung cancer cases and 730 (76·3%) of 957 lung cancer deaths occurring over 5 years (table 2). Less than half of this population (30 821 [14·2%]) was eligible by USPSTF-2013 criteria, which included 773 (48·8%) of 1585 lung cancer cases and 488 (51·0%) of 957 lung cancer deaths. NELSON criteria selected fewer individuals than USPSTF-2021 (64 851 [30·0%]) and fewer lung cancer cases (1082 [68·3%]) and deaths (674 [70·4%]). The median predicted 5-year risk of lung cancer death by LCDRAT for USPSTF-2013-eligible participants was 1·5% (IQR 1·0–2·4), and 0·8% (0·4–1·5) for each of USPSTF-2021 and NELSON (appendix 1 pp 6–7).In the National Lung Screening Trial, participants with a 5-year lung cancer death risk below 0·55% had no apparent absolute benefit from screening.[28] The percentage of participants in this category in our cohort data, who might also have no absolute benefit from screening, was 5·9% among those eligible by USPSTF-2013, 35·1% by USPSTF-2021, and 33·4% by NELSON criteria, although these percentages would differ in nationally representative data (appendix 1 pp 6–7).

For each prediction model, we identified risk thresholds that would screen the same number of individuals as each of the USPSTF-2021, USPSTF-2013, and NELSON criteria. Overall, this analysis showed that risk models identified slightly older participants (age difference ranging from 1 to 6 years) compared with categorical criteria (table 2). Considering screening efficiency to identify future lung cancer cases, we found that the $PLCO_{m2012}$, LCDRAT, LCRAT, Bach, HUNT, OWL, UCLD, and UCLI models classified similar numbers of future cases as eligible for screening, with OWL, LCDRAT, LCRAT, and HUNT showing slight advantages compared with the other models across the three criteria (table 2). For example, the USPSTF-2021 criteria screened 73 313 individuals and identified 1185 (74·8%) lung cancer cases over 5 years as screening-eligible. When screening the same number of individuals, the OWL, HUNT, LCDRAT, LCRAT, Bach, $PLCO_{m2012}$, UCLD, and UCLI models identified 77·6–79·1% of future cases as screening-eligible (number of cases ranging from 1230 to 1254), whereas LLP version 2 (1126 [71·0%]) and LLP version 3 (1115 [70·3%]) identified fewer cases. Results were similar for identifying future lung cancer deaths.

See Online for appendix 2

The screening-eligible populations identified by each model differed from the populations identified by categorical criteria. Figure 3 shows Venn diagrams comparing the participants and lung cancer cases selected by categorical criteria compared with the LCDRAT, HUNT, and $PLCO_{m2012}$ risk models, as examples of well-performing models developed in the USA (LCDRAT and $PLCO_{m2012}$) and Europe (HUNT). Taking USPSTF-2021 as an example (figure 3A), 49 093 participants were eligible by all of USPSTF-2021 and the LCDRAT, HUNT, and $PLCO_{m2012}$ models, and this group had high cumulative lung cancer incidence (2·1% over 5 years). The four groups of participants selected by only one approach ranged in size from 3070 to 8069 participants and had low lung cancer incidence (0·3–0·4%). Additional groups of participants were selected by two or three strategies, with incidence ranging from 0·4 to 1·2%. Analogous results are shown for USPSTF-2013 (figure 3B) and NELSON criteria (figure 3C).

When considering the individuals selected for screening by risk models but not categorical criteria, considering as examples the LCDRAT, HUNT, and $PLCO_{m2012}$ models, $PLCO_{m2012}$ selected more individuals with over 20 cigarettes smoked per day, over 30 pack-years, and a history of cancer, compared with the LCDRAT and HUNT models (appendix 1 pp 6–7). Both the HUNT and LCDRAT models selected more individuals with smoking duration over 40 years. $PLCO_{m2012}$ and LCDRAT selected more individuals with lower education levels, overweight or obesity, and a family history of lung cancer. The HUNT model selected more men and current smokers than the other models.

Discussion

As implementation of lung cancer screening progresses in Europe, evidence is needed to support the selection of strategies to define the target population. In this study, we compared the performance of ten models for predicting

	Threshold to screen the same number of participants as categorical criteria	Age, years (median [IQR])	Population selected	Lung cancer cases eligible for screening over 5 years	Lung cancer deaths eligible for screening over 5 years
Total	..	55 (50–60)	216 387	1585 (100·0%)	957 (100·0%)
USPSTF-2021 (aged 50–80 years, ≥20 pack-years, quit ≤15 years)					
USPSTF-2021	..	57 (53–61)	73 313 (33·9%)	1185 (74·8%)	730 (76·3%)
OWL	0·57%	58 (54–62)	73 312	1254 (79·1%)	767 (80·1%)
HUNT	0·71% (6-year time horizon)	58 (54–62)	73 313	1252 (79·0%)	774 (80·9%)
LCDRAT	0·45%	59 (55–63)	73 312	1250 (78·9%)	762 (79·6%)
LCRAT	0·85%	59 (55–63)	73 313	1241 (78·3%)	769 (80·4%)
UCLI	0·69%	59 (56–63)	73 316	1240 (78·2%)	757 (79·1%)
Bach	0·68%	59 (55–63)	73 340	1237 (78·0%)	757 (79·1%)
UCLD	0·45%	59 (55–63)	73 320	1233 (77·8%)	752 (78·6%)
$PLCO_{m2012}$	0·75% (6-year time horizon)	58 (54–62)	73 313	1230 (77·6%)	751 (78·5%)
LLP version 2	0·86%	61 (59–64)	73 882	1126 (71·0%)	683 (71·4%)
LLP version 3	0·46%	61 (59–64)	73 288	1115 (70·3%)	674 (70·4%)
USPSTF-2013 (aged 55–80 years, ≥30 pack-years, quit ≤15 years)					
USPSTF-2013	..	60 (57–63)	30 821 (14·2%)	773 (48·8%)	488 (51·0%)
LCDRAT	1·06%	61 (58–64)	30 820	864 (54·5%)	568 (59·4%)
LCRAT	1·86%	61 (58–64)	30 820	854 (53·9%)	552 (57·7%)
HUNT	1·41% (6-year time horizon)	61 (57–64)	30 820	848 (53·5%)	530 (55·4%)
OWL	1·27%	61 (57–64)	30 821	836 (52·7%)	525 (54·9%)
$PLCO_{m2012}$	1·72% (6-year time horizon)	61 (57–64)	30 821	832 (52·5%)	533 (55·7%)
Bach	1·63%	61 (58–64)	30 821	830 (52·4%)	523 (54·6%)
UCLI	1·52%	62 (59–64)	30 820	816 (51·5%)	514 (53·7%)
UCLD	0·92%	62 (59–64)	30 821	816 (51·5%)	512 (53·5%)
LLP version 3	1·33%	62 (60–65)	31 507	753 (47·5%)	473 (49·4%)
LLP version 2	2·51%	62 (60–65)	29 844	747 (47·1%)	466 (48·7%)
NELSON (aged 50–74 years, >15 cigarettes per day for >25 years or >10 cigarettes per day for >30 years, quit ≤10 years)					
NELSON	..	56 (52–60)	64 851 (30·0%)	1082 (68·3%)	674 (70·4%)
HUNT	0·80% (6-year time horizon)	58 (54–62)	64 851	1205 (76·0%)	746 (78·0%)
OWL	0·66%	59 (54–63)	64 851	1195 (75·4%)	736 (76·9%)
LCDRAT	0·52%	60 (56–63)	64 850	1191 (75·1%)	737 (77·0%)
LCRAT	0·98%	59 (55–63)	64 851	1185 (74·8%)	738 (77·1%)
Bach	0·80%	59 (56–63)	64 851	1185 (74·8%)	723 (75·5%)
$PLCO_{m2012}$	0·88% (6-year time horizon)	58 (54–62)	64 851	1177 (74·3%)	723 (75·5%)
UCLD	0·51%	59 (55–63)	64 850	1175 (74·1%)	714 (74·6%)
UCLI	0·79%	60 (56–63)	64 851	1167 (73·6%)	717 (74·9%)
LLP version 2	0·98%	62 (59–64)	65 090	1090 (68·8%)	662 (69·2%)
LLP version 3	0·52%	62 (59–64)	64 311	1056 (66·6%)	636 (66·5%)

Data are %, n (%), or n unless stated otherwise. For each categorical criterion (USPSTF-2021, USPSTF-2013, NELSON), a threshold for each model was selected so that the model would screen the same number of participants as the categorical criterion. Individuals whose predicted risk was higher than the threshold were considered screening-eligible. Then, within each eligible population, a higher number of lung cancer cases or deaths suggests a better performing model and more efficient screening. A single imputed dataset from the ATBC, CONSTANCES, and EPIC cohorts was used for the analyses in this table. The HUNT cohort was excluded from this analysis because it was used to develop the HUNT model. Risk was calculated over 5 years unless stated otherwise. ATBC=Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study. EPIC=European Prospective Investigation into Cancer and Nutrition. HUNT=Nord-Trøndelag Health Study. LCDRAT=Lung Cancer Death Risk Assessment Tool. LCRAT=Lung Cancer Risk Assessment Tool. NELSON=Nederlands-Leuvens Longkanker Screenings Onderzoek. LLP version 2=Liverpool Lung Project version 2. LLP version 3=Liverpool Lung Project version 3. OWL=Optimized Early Warning Model for Lung Cancer Risk. $PLCO_{m2012}$= Prostate, Lung, Colorectal, and Ovarian 2012 model. UCLD=University College London—Death. UCLI=University College London—Incidence. USPSTF-2021=US Preventive Services Task Force 2021. USPSTF-2013=US Preventive Services Task Force 2013.

***Table 2*: Performance of risk prediction models for defining lung cancer screening eligibility among current and former smokers aged 45–80 years, in pooled data from cohorts in nine European countries**

lung cancer incidence and mortality in nine European countries. Most models were reasonably calibrated in most countries, although LLP version 2 often overestimated risk. The $PLCO_{m2012}$, LCDRAT, LCRAT, Bach, HUNT, OWL, UCLD, and UCLI models showed similar discrimination in most countries, whereas the LLP version 2 and LLP version 3 models showed lower discrimination. When we defined risk thresholds to select the same number of

individuals as would be screened under three currently accepted categorical criteria, most models identified more future lung cancers than the categorical criteria.

Our study provides the most comprehensive assessment thus far of the performance of different lung cancer risk models in Europe. This is important for several reasons. The European Commission hesitated to promote national lung cancer screening in the absence of additional evidence on the effectiveness of screening strategies.[6] Major European research initiatives have attempted to identify populations at high risk by use of risk models instead of categorical criteria,[29] but there is no clear consensus on optimal screening eligibility criteria in Europe. There is substantial variation in eligibility criteria across lung cancer screening initiatives in Europe, with up to 15 years' difference in inclusion age and a 20 pack-year difference in smoking history.[30] Our study provides a direct comparison of three commonly used categorical screening eligibility criteria, together with ten risk prediction models. Owing to its reliance on data from prospective cohorts, we emphasise that our study provides a suitable setting for benchmarking the performance of different screening criteria, with the limitation that it cannot provide nationally representative results on the characteristics of the population selected for screening by different strategies.

Before this study, an analysis of two US cohorts found that the Bach, $PLCO_{m2012}$, LCRAT, and LCDRAT models outperformed LLP version 2 in both calibration and discrimination.[7] In three UK cohorts, LCDRAT, LCRAT, and Bach showed good discrimination with AUCs exceeding 0·80.[9] The reduced HUNT model, which excludes cough and secondhand smoke exposure from the predictors, has only been externally validated among 4051 heavy smokers from the Danish Lung Cancer Screening Trial.[12] In line with previous research, we found that LLP version 2 and LLP version 3 had lower discrimination than other models, probably because these models use categories instead of continuous parameters to model effects for age and smoking. However, we found only modest differences in discrimination between the Bach, $PLCO_{m2012}$, LCRAT, LCDRAT, HUNT, OWL, UCLD, and UCLI models, all of which include detailed information on age, smoking history, and other factors. The CanPredict (lung) model was developed and validated in the UK population, but we were unable to evaluate the performance of this model because the published parameters are not sufficient to implement the risk calculation.[31]

We found the discrimination of the HUNT model to be similar to the US models, although it underestimated lung cancer risk in some countries. This is probably owing to the use of a relatively healthy Norwegian population to develop the HUNT model.[32] Consistent with this, the remaining nine models overestimated risk in the HUNT cohort. The US models showed good risk-discriminatory performance, although LCRAT and LCDRAT overestimated risk in some countries. There are

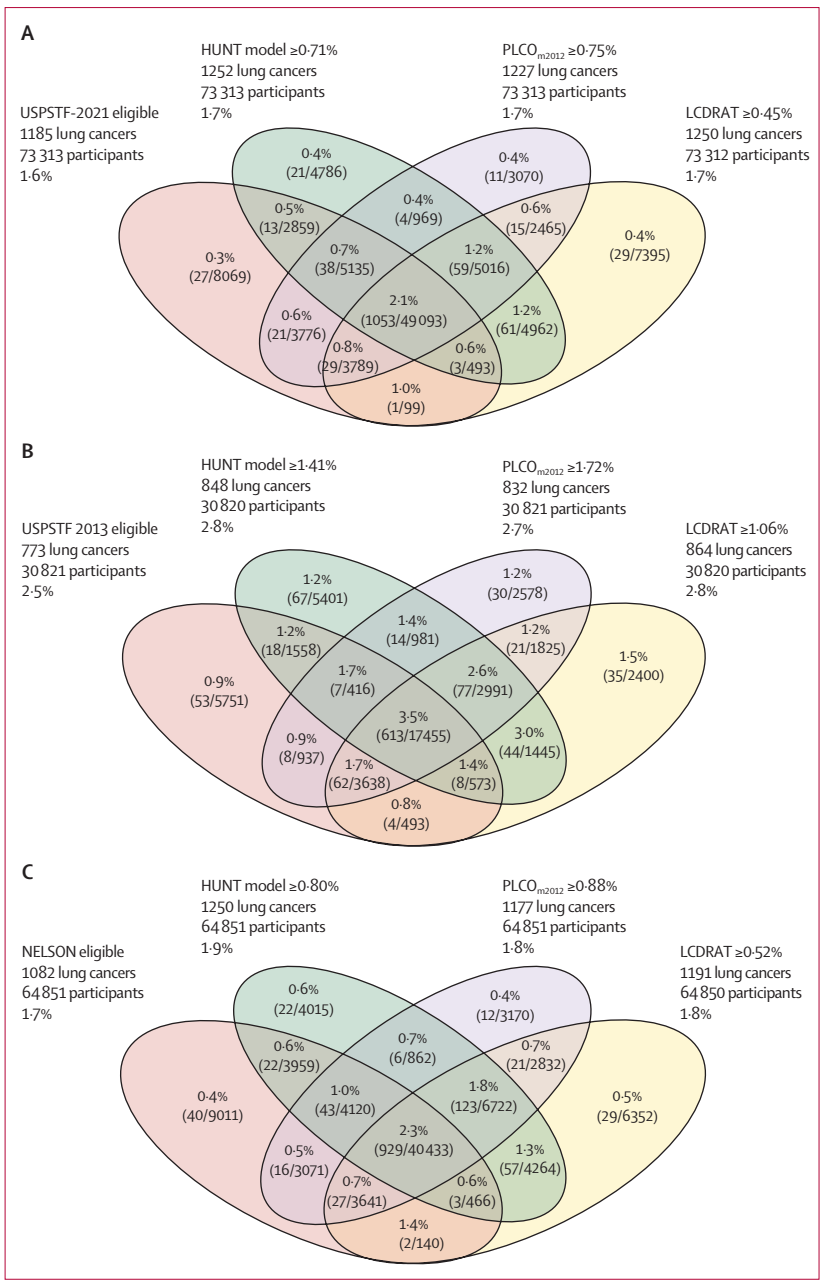

Figure 3: **Description of individuals and lung cancer cases eligible for lung cancer screening based on three risk prediction models as compared with the USPSTF-2021, USPSTF-2013, and NELSON criteria, in pooled data from cohorts from nine European countries**

Each diagram compares the participants selected by a categorical strategy (USPSTF-2021 [A], USPSTF-2013 [B], and NELSON [C]) with the groups of participants selected by the HUNT, $PLCO_{m2012}$, and LCDRAT risk models, when a threshold for each model is identified to select the same number of participants as the categorical strategy. Each cell shows the cumulative incidence of lung cancer over five years (cases/population). USPSTF 2021, age 50–80 years, at least 20 pack-years, quit ≤15 years. USPSTF 2013, age 55–80 years, at least 30 pack-years, quit ≤15 years. NELSON, age 50–74 years, >15 cigarettes a day for >25 years or >10 cigarettes per day for >30 years, quit ≤10 years. A single imputed dataset from the ATBC, CONSTANCES, and EPIC cohorts was used for the analyses in this figure. ATBC=Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study. EPIC=European Prospective Investigation into Cancer and Nutrition. HUNT=Nord-Trøndelag Health Study. LCDRAT=Lung Cancer Death Risk Assessment Tool. NELSON=Nederlands–Leuvens Longkanker Screenings Onderzoek. $PLCO_{m2012}$=Prostate, Lung, Colorectal, and Ovarian 2012 model. USPSTF-2021=US Preventive Services Task Force 2021. USPSTF-2013=US Preventive Services Task Force 2013.

multiple potential reasons for this that relate to US–Europe population differences or statistical aspects of these models, such as their included component of competing mortality. LLP version 3 is the updated version of LLP version 2, which was calibrated to national lung cancer data from England for 2017,[15] and it showed better calibration than LLP version 2 in most countries in this study. The LLP models identified fewer cases than categorical criteria. These models rely on smoking duration in wide categories (ie, 20–39 years, 40–59 years, etc) and do not account for smoking intensity (ie, cigarettes per day). The LLP models were therefore unable to identify most cases occurring among younger people (aged 50–54 years) whose high risk was due to high smoking intensity rather than long duration. To our knowledge, this is the first study to independently validate the OWL, UCLD, and UCLI ensemble machine learning-based models.[13,14] They showed similar calibration and discrimination to other well-performing models. Ultimately, differences in the well-performing models were small, and the thoughtful use of any of these models to define screening eligibility in European populations is likely to be reasonable.

When analysing data from these population cohorts, with a low representation of older people, the risk models identified more future lung cancer cases and deaths compared with categorical criteria, while selecting a slightly older population on average. It is likely that the relatively small improvement in sensitivity with risk models versus categorical criteria would be larger in a representative population, and that the age difference for screening-eligible individuals between categorical and risk-based criteria would widen if older people were fully represented. This question can be addressed by simulating the nationally representative population with detailed lung cancer risk factors.[33] We found that even when models had similar overall performance, they often selected populations with somewhat different characteristics. For example, $PLCO_{m2012}$ and LCDRAT were more likely to select individuals with lower socioeconomic status, among whom participation in lung cancer screening should be encouraged to increase screening effectiveness and reduce disparities.[34] Whereas risk models are already used to establish screening eligibility in some countries, such as in the UK,[35] financial and practical barriers have been reported elsewhere.[26,36] It is therefore important to evaluate the feasibility of implementing risk models in each country or health system.

By using a large, pooled European cohort database and excluding data used to develop the risk models where applicable, we provide a fully external and direct comparison of the major risk prediction models that might be considered for use in defining eligibility for lung cancer screening in Europe. However, our study has several limitations. First, there are some limitations to the generalisability of our results due to the nature of population cohort studies. The results of our study cannot be assumed to be nationally representative for each country. Differences across cohorts in the variance of predictors (eg, age, smoking) produce differences in discrimination that are artificial. Across the cohorts, individuals older than 70 years are absent or under-represented, which might decrease the discrimination of the risk models and reduce differences between categorical and risk-based criteria, as discussed in the previous paragraph. Some observations in specific cohorts, such as underprediction of risk in ATBC-Finland, overprediction of risk in HUNT-Norway and EPIC-Netherlands, and overprediction by mortality models in CONSTANCES-France, are probably due to specific aspects of these cohorts or their data collection, rather than the prediction models themselves. Second, smoking patterns have changed across Europe since the time when many cohort participants were enrolled.[8] Third, in CONSTANCES-France, 57% of participants have been followed up for less than 5 years, which might have caused some models to artificially overestimate risk in CONSTANCES, including the mortality models. Fourth, there was a substantial amount of missing data for some of the risk factors (often 100% missingness; for example family history of lung cancer, COPD or emphysema, asbestos exposure, daily cough, secondhand smoking, and cigarettes smoked per day were not collected in EPIC-Sweden). We handled this using the best possible method to avoid bias (multiple imputation).

In conclusion, our study evaluated the performance of ten lung cancer risk prediction models for selecting individuals for screening in nine European countries. We found that eight models performed well, with only minor differences among them (the $PLCO_{m2012}$, LCDRAT, LCRAT, Bach, HUNT, OWL, UCLD, and UCLI models), whereas the LLP version 2 and LLP version 3 models had lower risk-discriminatory performance. Our results indicate that well-performing risk models can identify more future lung cancer cases and deaths as eligible for low-dose CT screening than existing categorical screening criteria, without screening more people. European countries proceeding with lung cancer screening might consider testing the implementation of a well-performing risk prediction model to improve screening efficiency.

Contributors

MJ and HAR contributed to the conceptualisation, methodology, and supervision. XF, KA, and FG contributed to the data curation. XF, PG, and KA contributed to the data analysis. XF, PG, KA, FG, MJ, and HAR had full access to all study data. Access for all authors to the raw data was not feasible due to legal restrictions. All authors had access to the data presented in this study. XF, PG, and KA have directly accessed and verified the underlying data reported in the manuscript. XF, PG, and HAR drafted the manuscript. RK, RV, GSD, CB, SMC-Y, DA, SJW, MG, MZ, CR, AL, and AHS contributed data. All authors read and approved the final version of the manuscript.

Declaration of interests

We declare no competing interests. Where authors are identified as personnel of the International Agency for Research on Cancer or WHO, the authors alone are responsible for the views expressed in this Article and they do not necessarily represent the decisions, policy, or views of those organisations.

Data sharing

Access to data from the Lung Cancer Cohort Consortium (LC3) is governed by the LC3 Access Policy, which is available at the following link: https://www.iarc.who.int/wp-content/uploads/2021/12/LC3_Access_Policy.pdf. Interested investigators are encouraged to contact Mattias Johansson or Hilary A Robbins to propose a project using LC3 data. Codes and information for the 10 lung cancer prediction models are available on Github: https://github.com/IET-IARC/LungCancerModels.

Acknowledgments

This study was done in the context of the Lung Cancer Cohort Consortium (LC3) project which was approved by the Ethics Committee of the International Agency for Research on Cancer (approval number: 11-13 and 22-31). All cohort participants provided informed consent at enrolment. This study was supported by the US National Cancer Institute (R03CA245979 and U19CA203654), l'Institut National du Cancer (French National Cancer Institute, DEPREV 2020-126), Cancer Research UK (C18281/A19169), and the North West Lung Centre Charity (PhD studentship for PG).

References

1 Aberle DR, Adams AM, Berg CD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med* 2011; **365:** 395–409.

2 Paci E, Puliti D, Lopes Pegna A, et al. Mortality, survival and incidence rates in the ITALUNG randomised lung cancer screening trial. *Thorax* 2017; **72:** 825–31.

3 Becker N, Motsch E, Trotter A, et al. Lung cancer mortality reduction by LDCT screening—results from the randomized German LUSI trial. *Int J Cancer* 2020; **146:** 1503–13.

4 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:** 503–13.

5 Wait S, Alvarez-Rosete A, Osama T, et al. implementing lung cancer screening in Europe: taking a systems approach. *JTO Clin Res Rep* 2022; **3:** 100329.

6 European Commission: DG Health and Food Safety. Proposal for a Council Recommendation on strengthening prevention through early detection: a new EU approach on cancer screening replacing Council Recommendation 2003/878/EC, 2022. https://health.ec.europa.eu/publications/proposal-council-recommendation-cr-strengthening-prevention-through-early-detection-new-approach_en (accessed May 8, 2023).

7 Katki HA, Kovalchik SA, Petito LC, et al. Implications of nine risk prediction models for selecting ever-smokers for computed tomography lung cancer screening. *Ann Intern Med* 2018; **169:** 10–19.

8 Reitsma MB, Kendrick PJ, Ababneh E, et al. Spatial, temporal, and demographic patterns in prevalence of smoking tobacco use and attributable disease burden in 204 countries and territories, 1990–2019: a systematic analysis from the Global Burden of Disease Study 2019. *Lancet* 2021; **397:** 2337–60.

9 Robbins HA, Alcala K, Swerdlow AJ, et al. Comparative performance of lung cancer risk models to define lung screening eligibility in the United Kingdom. *Br J Cancer* 2021; **124:** 2026–34.

10 Li K, Hüsing A, Sookthai D, et al. Selecting high-risk individuals for lung cancer screening: a prospective evaluation of existing risk models and eligibility criteria in the German EPIC cohort. *Cancer Prev Res (Phila)* 2015; **8:** 777–85.

11 Cronin KA, Gail MH, Zou Z, Bach PB, Virtamo J, Albanes D. Validation of a model of lung cancer risk prediction among smokers. *J Natl Cancer Inst* 2006; **98:** 637–40.

12 Røe OD, Markaki M, Tsamardinos I, et al. 'Reduced' HUNT model outperforms NLST and NELSON study criteria in predicting lung cancer in the Danish screening trial. *BMJ Open Respir Res* 2019; **6:** e000512.

13 Callender T, Imrie F, Cebere B, et al. Assessing eligibility for lung cancer screening using parsimonious ensemble machine learning models: a development and validation study. *PLoS Med* 2023; **20:** e1004287.

14 Pan Z, Zhang R, Shen S, et al. OWL: an optimized and independently validated machine learning prediction model for lung cancer screening based on the UK Biobank, PLCO, and NLST populations. *EBioMedicine* 2023; **88:** 104443.

15 Field JK, Vulkan D, Davies MPA, Duffy SW, Gabe R. Liverpool Lung Project lung cancer risk stratification model: calibration and prospective validation. *Thorax* 2021; **76:** 161–68.

16 Robbins HA, Alcala K, Moez EK, et al. Design and methodological considerations for biomarker discovery and validation in the Integrative Analysis of Lung Cancer Etiology and Risk (INTEGRAL) Program. *Ann Epidemiol* 2023; **77:** 1–12.

17 Riboli E, Hunt KJ, Slimani N, et al. European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. *Public Health Nutr* 2002; **5:** 1113–24.

18 Zins M, Goldberg M. The French CONSTANCES population-based cohort: design, inclusion and follow-up. *Eur J Epidemiol* 2015; **30:** 1317–28.

19 Cohorte CONSTANCES. https://doi.org/10.13143/inserm_constances (accessed March 29, 2024).

20 Bach PB, Kattan MW, Thornquist MD, et al. Variations in lung cancer risk among smokers. *J Natl Cancer Inst* 2003; **95:** 470–78.

21 Tammemägi MC, Katki HA, Hocking WG, et al. Selection criteria for lung-cancer screening. *N Engl J Med* 2013; **368:** 728–36.

22 Katki HA, Kovalchik SA, Berg CD, Cheung LC, Chaturvedi AK. Development and validation of risk models to select ever-smokers for CT lung cancer screening. *JAMA* 2016; **315:** 2300–11.

23 Markaki M, Tsamardinos I, Langhammer A, Lagani V, Hveem K, Røe OD. A validated clinical risk prediction model for lung cancer in smokers of all ages and exposure types: a HUNT study. *EBioMedicine* 2018; **31:** 36–46.

24 Cheung L, Katki H. lcmodels (R package). 2022. NCI Division of Cancer Epidemiology and Genetics: tools and resources. https://dceg.cancer.gov/tools/risk-assessment/lcmodels (accessed March 24, 2022).

25 Rubin DB. Multiple imputation for nonresponse in surveys. New York: John Wiley & Sons, 2004.

26 Krist AH, Davidson KW, Mangione CM, et al. Screening for lung cancer: US Preventive Services Task Force recommendation statement. *JAMA* 2021; **325:** 962–70.

27 Moyer VA. Screening for lung cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2014; **160:** 330–38.

28 Kovalchik SA, Tammemagi M, Berg CD, et al. Targeting of low-dose CT screening according to the risk of lung-cancer death. *N Engl J Med* 2013; **369:** 245–54.

29 4-IN THE LUNG RUN: Towards INdividually tailored INvitations screening INtervals and INtegrated comorbidity reducing strategies in lung cancer screening. https://www.erasmusmc-rdo.nl/project/towards-individually-tailored-invitations-screening-intervals-and-integrated-co-morbidity-reducing-strategies-in-lung-cancer-screening/ (accessed Sept 30, 2022).

30 Lung Cancer Policy Network. Interactive map of lung cancer screening. 2022. https://www.lungcancerpolicynetwork.com/interactive-map-of-lung-cancer-screening/ (accessed Sept 30, 2022).

31 Liao W, Coupland CAC, Burchardt J, et al. Predicting the future risk of lung cancer: development, and internal and external validation of the CanPredict (lung) model in 19·67 million people and evaluation of model performance against seven other risk prediction models. *Lancet Respir Med* 2023; **11:** 685–97.

32 European Commission. State of Health in the EU. Country Health Profile 2021, Norway. 2021. https://health.ec.europa.eu/system/files/2021-12/2021_chp_no_english.pdf (accessed June 14, 2023).

33 Miranda-Filho A, Charvat H, Bray F, et al. A modeling analysis to compare eligibility strategies for lung cancer screening in Brazil. *EClinicalMedicine* 2021; **42:** 101176.

34 Baldwin DR, Brain K, Quaife S. Participation in lung cancer screening. *Transl Lung Cancer Res* 2021; **10:** 1091–98.

35 NHS England—NHS Cancer Programme. Targeted screening for lung cancer with low radiation dose computed tomography. Standard protocol prepared for the NHS England Targeted Lung Health Checks Programme, version 2. 2022. https://www.england.nhs.uk/publication/targeted-screening-for-lung-cancer/ (accessed May 8, 2023).

36 Van Hal G, Diab Garcia P. Lung cancer screening: targeting the hard to reach—a review. *Transl Lung Cancer Res* 2021; **10:** 2309–22.