Published in final edited form as:

Ann Intern Med. 2012 August 21; 157(4): 242–250. doi:10.7326/0003-4819-157-4-201208210-00004.

Predictive Accuracy of the Liverpool Lung Project Risk Model for Stratifying Patients for Computed Tomography Screening for Lung Cancer:

A Case–Control and Cohort Validation Study

Olaide Y. Raji, PhD, Stephen W. Duffy, MSc, Olorunshola F. Agbaje, PhD, Stuart G. Baker, ScD, David C. Christiani, MD, MPH, Adrian Cassidy, PhD, and John K. Field, PhD, FRCPath Roy Castle Lung Cancer Research Programme, The University of Liverpool, Liverpool, United Kingdom; Wolfson Institute of Preventive Medicine, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, King's College London, and Guy's Hospital, London, United Kingdom; National Cancer Institute, Bethesda, Maryland; and Harvard School of Public Health, Boston, Massachusetts.

Abstract

Background—External validation of existing lung cancer risk prediction models is limited. Using such models in clinical practice to guide the referral of patients for computed tomography

© 2012 American College of Physicians

Requests for Single Reprints: John K. Field, PhD, FRCPath, Roy Castle Lung Cancer Research Programme, The University of Liverpool Cancer Research Centre, Institute of Translational Medicine, The University of Liverpool, Liverpool L3 9TA, United Kingdom; J.K.Field@liv.ac.uk.

Current Author Addresses: Dr. Raji and Professor Field: Roy Castle Lung Cancer Research Programme, The University of Liverpool Cancer Research Centre, Institute of Translational Medicine, The University of Liverpool, Liverpool L3 9TA, United Kingdom.

Professor Duffy: Wolfson Institute of Preventive Medicine, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, Charterhouse Square, London EC1M 6BQ, United Kingdom.

Dr. Agbaje: Division of Cancer Studies, Cancer Epidemiology Unit, King's College London, School of Medicine, Academic Oncology, 3rd Floor, Bermondsey Wing, Guy's Hospital, London SE1 9RT, United Kingdom. Dr. Baker: National Cancer Institute, EPN 3131, 6130 Executive Boulevard MSC, Bethesda, MD 20892.

Dr. Christiani: Harvard School of Public Health, 665 Huntington Avenue, Boston, MA 02115.

Dr. Cassidy: GlaxoSmithKline Vaccines, Parc de la Noire Epine, Rue Fleming 20, 1300 Wavre, Belgium.

Potential Conflicts of Interest: Disclosures can be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M11-1994.

Reproducible Research Statement: *Study protocol:* Available from Professor Field (J.K.Field@liv.ac.uk). *Statistical code:* Available from Dr. Raji (O.Y.Raji@liv.ac.uk). Relative utility curves are available from Dr. Baker (sb16i@nih.gov). *Data set:* LLCC and LLPC data are available from Professor Field (J.K.Field@liv.ac.uk) on completion of a data transfer agreement. Harvard and EUELC data need separate approval for their release.

Author Contributions: Conception and design: S.W. Duffy, A. Cassidy, J.K. Field.

Analysis and interpretation of the data: O.Y. Raji, S.W. Duffy, O.F.

Agbaje, S.G. Baker, D.C. Christiani, A. Cassidy, J.K. Field.

Drafting of the article: O.Y. Raji, O.F. Agbaje, A. Cassidy, J.K. Field.

Critical revision of the article for important intellectual content: O.Y.

Raji, S.G. Baker, D.C. Christiani, A. Cassidy, J.K. Field.

Final approval of the article: O.Y. Raji, S.W. Duffy, O.F. Agbaje, D.C.

Christiani, A. Cassidy, J.K. Field.

Provision of study materials or patients: D.C. Christiani, J.K. Field.

Statistical expertise: S.W. Duffy, S.G. Baker.

Obtaining of funding: D.C. Christiani, J.K. Field.

Administrative, technical, or logistic support: D.C. Christiani, J.K. Field.

Collection and assembly of data: O.Y. Raji, A. Cassidy, J.K. Field.

(CT) screening for lung cancer depends on external validation and evidence of predicted clinical benefit.

Objective—To evaluate the discrimination of the Liverpool Lung Project (LLP) risk model and demonstrate its predicted benefit for stratifying patients for CT screening by using data from 3 independent studies from Europe and North America.

Design—Case-control and prospective cohort study.

Setting—Europe and North America.

Patients—Participants in the European Early Lung Cancer (EUELC) and Harvard case-control studies and the LLP population-based prospective cohort (LLPC) study.

Measurements—5-year absolute risks for lung cancer predicted by the LLP model.

Results—The LLP risk model had good discrimination in both the Harvard (area under the receiver-operating characteristic curve [AUC], 0.76 [95% CI, 0.75 to 0.78]) and the LLPC (AUC, 0.82 [CI, 0.80 to 0.85]) studies and modest discrimination in the EUELC (AUC, 0.67 [CI, 0.64 to 0.69]) study. The decision utility analysis, which incorporates the harms and benefit of using a risk model to make clinical decisions, indicates that the LLP risk model performed better than smoking duration or family history alone in stratifying high-risk patients for lung cancer CT screening.

Limitations—The model cannot assess whether including other risk factors, such as lung function or genetic markers, would improve accuracy. Lack of information on asbestos exposure in the LLPC limited the ability to validate the complete LLP risk model.

Conclusion—Validation of the LLP risk model in 3 independent external data sets demonstrated good discrimination and evidence of predicted benefits for stratifying patients for lung cancer CT screening. Further studies are needed to prospectively evaluate model performance and evaluate the optimal population risk thresholds for initiating lung cancer screening.

Primary Funding Source—Roy Castle Lung Cancer Foundation.

Lung cancer is the leading cause of cancer mortality worldwide, with more than 1 million deaths each year (1). The disease is usually diagnosed at an advanced stage, when surgical resection or other treatment options are less effective, thus leading to poor survival (2). The results from the NLST (National Lung Screening Trial), sponsored by the National Cancer Institute, showed a 20% decrease in lung cancer deaths and a 6% decrease in all-cause mortality when smokers were screened annually for 3 years with low-dose spiral computed tomography (CT) compared with standard chest radiography (3). The ongoing evaluations of CT screening have indicated the need to identify persons at high risk for the disease to maximize the benefit-harm ratio (4–6). Although this approach has been endorsed by the International Association for the Study of Lung Cancer in its recent report (7), no tools adequately identify which patients should be targeted to maximize the screening benefit.

Many risk models have been developed to predict individual risk for lung cancer within a specified period by using patient characteristics and epidemiologic, social, and clinical risk factors (8). These include risk models by Bach and colleagues (9), Spitz and colleagues (10), Tammemagi and colleagues (11), and the Liverpool Lung Project (LLP) (12). The LLP risk model, developed from the LLP case-control (LLCC) study, provides a single unified model for smokers (current and former) and nonsmokers, whereas the Bach model was developed for predicting risk only in smokers and the Spitz model requires 3 separate models for predicting risk in current, former, or nonsmokers. In addition, the LLP model accounts for important lung cancer risk factors besides age, sex, and smoking duration, including history of pneumonia, history of nonlung cancer, asbestos exposure, and family

history. This makes it simpler than Tammemagi and colleagues' model, which includes many smoking-related variables that may be difficult to obtain in a clinical setting.

A clinically useful risk model needs to perform well in independent data sets and be generalizable to external populations (13–16). Traditional statistical measures for evaluating risk models are the area under the receiver-operating characteristic curve (AUC), which quantifies the model's ability to discriminate between patients with and those without disease, and the calibration accuracy, which compares the predicted and observed disease probabilities (13, 16, 17). A key limitation of these predictive measures is that they focus purely on the mathematical accuracy of the model and do not assess predicted net benefit, which is based on both accuracy and tradeoffs between the harms and benefits of using the model for making clinical decisions (such as who should receive lung cancer CT screening).

Two recently proposed decision-theoretical methods that put risk model assessment into clinical perspective are decision curve analysis and the relative utility curve method (18–21). These methods explicitly incorporate the predicted harms and benefits of clinical decisions based on the model predictions at a given risk threshold. A risk threshold is the absolute risk for lung cancer at which one might choose to initiate screening. A person whose risk, as predicted by the model, is above the threshold would be stratified as high-risk and selected for screening. The decision curve and relative utility curve estimate the predicted net benefit of the model across all possible risk thresholds, making it easier to evaluate the effect of various risk thresholds.

Our study presents the validation of the LLP risk model in 2 independent case-control studies and a population-based prospective cohort study by assessing its discriminative performance. In addition, we evaluate the potential clinical effect of using the model for making decisions about lung cancer CT screening.

Methods

Study Population

We used data from the EUELC (European Early Lung Cancer) case-control study, the Harvard case-control study, and the LLP population-based prospective cohort (LLPC) study. The EUELC study (22) is an international, collaborative case-control study of early-stage lung cancer conducted in 8 European countries between 2002 and 2006. The 585 case patients had histologically or cytologically confirmed non-small cell lung cancer with surgically resected primary tumors and were matched with control participants by age, sex, and study center. The 1238 control participants were recruited from hospitals or from the population registers of general practitioners in the same area as the case patients. The Harvard case-control study (23) is a hospital-based study of cases of non-small cell lung cancer diagnosed at Massachusetts General Hospital, Boston, Massachusetts, since 1992. A total of 2922 of 3298 participants were used (1738 patients with histologically confirmed lung cancer and 1184 control participants). Control participants were family members and friends of case patients or persons attending the same hospital for other diseases.

The LLPC (24) comprised persons aged 40 to 79 years who were randomly selected from the Liverpool area (population cohort) or recruited from hospitals to which they came for health episodes other than lung cancer (hospital cohort). The cohort comprised 7652 participants recruited between 1998 and 2008 and followed annually for lung cancer and mortality outcomes through the Office for National Statistics, the North West Cancer Intelligence Service, and hospital case-note review. Participants in the LLPC and LLCC are mutually exclusive.

Ethical approval was obtained for each study from the local institutional review board or research ethics committee. All study participants provided written informed consent.

Data Collection and Extraction of Risk Factors

Each study used a standardized questionnaire to collect self-reported information on demographic and socioeconomic characteristics, medical history, family history of cancer, history of tobacco consumption (age of starting and quitting smoking and number of cigarettes per day), and lifetime occupational history. Information on the 5 risk factors in the LLP risk model (smoking duration, history of pneumonia, history of cancer, family history of cancer, and asbestos exposure) was extracted from the questionnaire. Smoking duration was measured in years (never smoked or smoked for <20, 21 to 40, 41 to 60, or >60 years). Previous pneumonia or cancer (except melanoma) were each coded as “yes” or “no.” Family history of lung cancer consisted of age at onset in a first-degree relative (none, early [<60 years], or late [≥ 60 years]). Occupational exposure to asbestos was individually assessed in the LLCC by using the “expert assessment” method (25), but this was resource-intensive and was not performed in the LLPC. Therefore, asbestos exposure was missing for 97% of LLPC participants.

Statistical Analysis

The LLP risk model, developed from the LLCC study, was used to predict a person’s absolute 5-year risk for lung cancer. The model development process and internal validation have been published elsewhere (12). Table 1 presents the estimated coefficients, derived at the model development stage by using the LLCC data, together with the coefficients of a refitted model without asbestos exposure (Appendix Table 1, available at www.annals.org, lists the coefficients of models without each of the 5 risk factors). Smoking duration was found to be the most important predictor, followed by age of lung cancer onset in a first-degree family member regardless of asbestos exposure. In addition, the 2 models had similar internal discrimination. Risk was predicted for participants in the LLPC by using the model without asbestos exposure. The diversity of predicted risks for a smoker and nonsmoker of similar age are published elsewhere (12), and Appendix Table 2 (available at www.annals.org) shows the risk conferred by factors other than smoking duration.

The performance of the LLP risk model was assessed by measuring discriminative accuracy (13, 16, 17) and by decision curve (18, 19) and relative utility curve analysis (20, 21). Discrimination was assessed by using the AUC, which measures the ability of the model to separate persons who will develop lung cancer from those who will not. The AUC values range from 0.5 (no [chance] discrimination) to 1 (perfect discrimination).

Decision curves plot the predicted net benefit of the risk prediction model versus risk thresholds. Relative utility curves plot the predicted net benefit of the model relative to the predicted net benefit of a perfect prediction against risk thresholds. The predicted net benefit is the total true-positive classifications minus the total false-positive classifications, weighted by the odds of the risk threshold (18). The weighting converts a false-positive classification into the same units as the true-positive classification; thus, the net benefit is interpreted as the number of true-positive classifications adjusted for the detrimental effect of false-positive classifications. To put the net benefit in context, we compare the net benefit of the LLP model with that of other clinical alternatives, such as screening no one or screening everyone, and calculate the net gain, which is the increase in net benefit from the LLP model versus a screen-all strategy. We also compare the model-predicted net benefit with that achieved by using smoking duration and family history of cancer, the 2 strongest predictors of lung cancer risk.

Statistical analyses were done using STATA, version 10.0 (StataCorp, College Station, Texas); R, version 2.10.0 (R Foundation for Statistical Computing, Vienna, Austria); and Mathematica, version 8.0 (Wolfram Research, Champaign, Illinois).

Role of the Funding Source

This work was supported by the Roy Castle Lung Cancer Foundation and partly by the National Institute for Health Research Health Technology Assessment directorate, the American Cancer Society, and the National Cancer Institute of the National Institutes of Health. The funding sources had no role in the design, conduct, or reporting of this study or in the decision to submit the manuscript for publication.

Results

Distribution of Participant Characteristics

Table 2 shows the distributions of participants' risk factors in the 3 validation data sets. Most participants in the EUELC were men, regardless of case–control status. Different sex distributions were observed for case patients and control participants in the Harvard study. The distributions of age, smoking duration, family history, and asbestos exposure followed similar patterns in the 3 studies, particularly for case patients.

In the LLPC, 420 of 7652 participants (approximately 6% of the cohort) developed lung cancer over an average follow-up of 8 years (median, 7 years). The lung cancer rates were slightly higher in men than in women, higher in participants with a history of pneumonia than in those without, and approximately 3 times higher in persons with a history of cancer than in those without. Lung cancer rates also increased with greater age and longer smoking duration.

Performance of the LLP Model

Risk Distribution—Figure 1 shows the distributions of predicted absolute risk by disease status in the 3 data sets. In general, individual absolute risks were lower for control participants than for case patients. Most risks greater than 2.5% were predicted for patients with cancer, whereas about one half of disease-free patients had absolute risks less than 1%. The median 5-year predicted values were substantially lower for control participants than for case patients, indicating good separation of summary values for patients with and without cancer.

Discrimination—The LLP risk model had higher discriminative ability across the 3 data sets than using smoking duration or family history of lung cancer (Table 3). The model had modest discrimination in the EUELC data set (AUC, 0.67 [CI, 0.64 to 0.69]) and good discrimination in both the Harvard (AUC, 0.76 [95% CI, 0.75 to 0.78]) and LLPC (AUC, 0.82 [CI, 0.80 to 0.85]) data sets. The AUC for smoking duration, the strongest of the risk factors, was 0.63, 0.74, and 0.72 in the EUELC, Harvard, and LLPC data sets, respectively (data not shown). The LLP risk model had moderate overall calibration and improved accuracy at higher values of predicted risks (Appendix Figure 1, available at www.annals.org).

Assessment of Potential for Clinical Application

Table 3 shows the LLP model's sensitivity, specificity, and estimated net benefit at thresholds of 2.5%, 5%, and 10% predicted absolute risk (Appendix Table 3, available at www.annals.org, shows the number of patients stratified by risk threshold). A positive net gain indicates that the model had greater net benefit than a screen-all strategy. At a threshold

of 5% absolute risk, the model achieved a higher proportion of true-positive classifications than a screen-all strategy (2.3% higher for the LLPC data and 3% higher for the EUELC data) at the same proportion of false-positive classifications. Panels A to C in Figure 2 compare the net benefits of using the LLP risk model, the 2 strongest risk factors (smoking duration and family history), or the extreme strategies of screening everyone or no one. The LLP risk model had greater net benefit than all alternative strategies at thresholds of absolute risk, ranging from 3% to 15%.

Of note, the LLP risk model performs well relative to the strong predictor of smoking duration, which is most often used to stratify high-risk persons for lung cancer CT screening. For the LLPC data, the receiver-operating characteristic curve showed a moderate increase in discrimination over smoking when using the LLP risk model. For relevant risk thresholds at a probability of disease greater than 0.05, the relative utility curve showed moderately higher predicted net benefit, relative to a perfect prediction, for the LLP risk model than for smoking duration (Figure 2, D). For the EUELC and Harvard data, the increase in predicted net benefit for the LLP risk model versus smoking duration at the relevant high-risk thresholds was smaller (Appendix Figure 2, available at www.annals.org).

Discussion

Using an independent validation approach, we have demonstrated that the LLP risk model has a good ability to distinguish persons who will or will not develop lung cancer by using the predicted 5-year absolute risk. The model also seems to be reasonably well-calibrated at high predicted risks and performs better than smoking duration or family history as a tool for deciding which persons to screen for lung cancer. The LLP risk model also unifies smoking duration, other important risk factors for lung cancer, and incidence data from cancer registries, thereby combining the benefit of each to provide accurate and diverse predicted risks for smokers and nonsmokers.

Risk prediction models with good discrimination are potentially valuable for identifying high-risk persons in a disease-screening application (26–28). The discrimination attained by the LLP risk model compared well with those reported for existing risk models for other types of cancer or chronic diseases, such as the Gail model for breast cancer (29–31), the LAMBDA model for familial breast cancer (32), and the Framingham risk score for coronary heart disease (33, 34). The LLP model's modest discrimination in the EUELC may be due to the case selection, because only surgically treated patients with non-small cell lung cancer were included in this project. In addition, whereas the Harvard and LLPC data sets are from a single and defined urban population, the EUELC comprised patients from 8 different populations with varying levels of risks. Patient case-mix may affect the discrimination of risk models because of differences in patient populations, risk factor distributions, and predictor effects over time (35, 36).

Assessing the potential clinical application of a risk model is an important part of validation because it illustrates the potential benefits and harms based on the predicted absolute risk (37). The use of lung cancer risk prediction models to select high-risk patients for lung cancer CT screening is limited, possibly because no model has had rigorous independent testing to determine its effect on patient care (26). Most CT trials, including the NLST and NELSON (Nederlands Leuvens Longkanker Screenings Onderzoek) trial, have used smoking history and age to identify high-risk persons (3, 6, 38). This strategy assumes similar levels of risk for everyone in a specific age–smoking combination. Although this approach was considered appropriate when the NLST and NELSON trial were initiated, the need for risk models that predict a person's chance of developing lung cancer is now

recognized (28), so that fewer persons are identified for screening and more of those identified are found to have lung cancer.

An important aspect of our validation is the sensitivity analysis of performance across various risk thresholds provided by the decision and relative utility curves. These thresholds are the levels of absolute risk for lung cancer at which one might choose to initiate screening. They imply the relative seriousness of false- and true-positive diagnoses. For example, a risk threshold of 5% would mean that it was acceptable to misdiagnose cancer in up to 95 persons to correctly diagnose it in 5, which implies that not detecting cancer is 19 times worse than falsely diagnosing it ($95/5 = 19$). This 19:1 ratio can be used to weight the false-positive diagnosis when applying a prediction model to a cohort, assuming that a person whose predicted risk exceeds the threshold would be selected for screening. The decision and relative utility curve estimates use the error weightings implied by each risk threshold to calculate the net benefit of using that model compared with such alternative strategies as screening everyone or no one.

Identifying a single average risk threshold for a population is often difficult because of a lack of data on harms, benefits, and actual outcomes in a screened population. Unlike cardiovascular disease, for which a 10-year risk of 20% has been recommended to stratify patients as high-risk (39), no consensus is available in cancer screening (40). Retrospective analysis of data from the NLST and CT screening studies in Europe, together with prospective evaluation and follow-up of patients in the UKLS (United Kingdom Lung Screening) trial (41), may help to standardize the risk threshold at which to recommend population-based CT screening for lung cancer.

Our analysis has limitations. The lack of asbestos exposure data for most LLPC participants precludes using the full model to predict risk in the cohort. However, our results suggest no considerable loss of accuracy or performance. Our analysis also provides only predicted benefits; it does not compute the actual benefits or explicitly recommend a particular risk threshold because these would require actual CT screening trial data with outcome information (26). Finally, although we considered many risk factors at the developmental stage of the LLP risk model, data on spirometry measurement (using an objective measure of chronic obstructive pulmonary disease) and genetic markers (such as single nucleotide polymorphisms from recent genome-wide association studies [42, 43]) were not assessed. Future refinement of the model would enable assessment of whether these factors further improve accuracy and utility.

In conclusion, the LLP risk model is a simple model with few variables and may be useful for selecting high-risk patients for lung cancer CT screening. Further prospective evaluation of the model is required before it can be used as a clinical tool in primary care.

Acknowledgments

The authors thank all study participants, the EUELC Consortium, Dr. Andrew J. Vickers for his useful discussion and helpful comments during the statistical data analysis and preparation of the manuscript, and Professor Anne Field for reading the manuscript as a nonexpert clinician. For members of the EUELC Consortium, see the Appendix (available at www.annals.org).

Grant Support: By the Roy Castle Lung Cancer Foundation, the National Institute for Health Research Health Technology Assessment program, and the American Cancer Society, as well as grants CA74386, CA092824, and CA090578 from the National Cancer Institute, National Institutes of Health (Dr. Christiani).

References

1. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin*. 2005; 55:74–108. [PMID: 15761078]. [PubMed: 15761078]
2. Schiller JH, Harrington D, Belani CP, Langer C, Sandler A, Krook J, et al. Eastern Cooperative Oncology Group. Comparison of four chemotherapy regimens for advanced non-small-cell lung cancer. *N Engl J Med*. 2002; 346:92–98. [PMID: 11784875]. [PubMed: 11784875]
3. Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, Fagerstrom RM, et al. National Lung Screening Trial Research Team. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med*. 2011; 365:395–409. [PMID: 21714641]. [PubMed: 21714641]
4. van Klaveren RJ, Habbema JDF, Pedersen JH, de Koning HJ, Oudkerk M, Hoogsteden HC. Lung cancer screening by low-dose spiral computed tomography. *Eur Respir J*. 2001; 18:857–866. [PMID: 11757637]. [PubMed: 11757637]
5. Field JK, Duffy SW. Lung cancer screening: the way forward. *Br J Cancer*. 2008; 99:557–562. [PMID: 18665179]. [PubMed: 18665179]
6. Lopes, Pegna A.; Picozzi, G. Lung cancer screening update. *Curr Opin Pulm Med*. 2009; 15:327–333. [PMID: 19395971]. [PubMed: 19395971]
7. Field JK, Smith RA, Aberle DR, Oudkerk M, Baldwin DR, Yankelevitz D, et al. IASLC CTScreening Workshop 2011 Participants. International Association for the Study of Lung Cancer Computed Tomography Screening Workshop 2011 report. *J Thorac Oncol*. 2012; 7:10–19. [PMID: 22173661]. [PubMed: 22173661]
8. Cassidy A, Duffy SW, Myles JP, Liloglou T, Field JK. Lung cancer risk prediction: a tool for early detection. *Int J Cancer*. 2007; 120:1–6. [PMID: 17058200]. [PubMed: 17058200]
9. Bach PB, Kattan MW, Thornquist MD, Kris MG, Tate RC, Barnett MJ, et al. Variations in lung cancer risk among smokers. *J Natl Cancer Inst*. 2003; 95:470–478. [PMID: 12644540]. [PubMed: 12644540]
10. Spitz MR, Hong WK, Amos CI, Wu X, Schabath MB, Dong Q, et al. A risk model for prediction of lung cancer. *J Natl Cancer Inst*. 2007; 99:715–726. [PMID: 17470739]. [PubMed: 17470739]
11. Tammemagi CM, Pinsky PF, Caporaso NE, Kvale PA, Hocking WG, Church TR, et al. Lung cancer risk prediction: Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial models and validation. *J Natl Cancer Inst*. 2011; 103:1058–1068. [PMID: 21606442]. [PubMed: 21606442]
12. Cassidy A, Myles JP, van Tongeren M, Page RD, Liloglou T, Duffy SW, et al. The LLP risk model: an individual risk prediction model for lung cancer. *Br J Cancer*. 2008; 98:270–276. [PMID: 18087271]. [PubMed: 18087271]
13. Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. *BMJ*. 2009; 338:b606. [PMID: 19502216]. [PubMed: 19502216]
14. Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. *BMJ*. 2009; 338:b605. [PMID: 19477892]. [PubMed: 19477892]
15. Efron B. How biased is the apparent error rate of a prediction rule? *J Am Stat Assoc*. 1986; 81:461–470.
16. Freedman AN, Seminara D, Gail MH, Hartge P, Colditz GA, Ballard-Barbash R, et al. Cancer risk prediction models: a workshop on development, evaluation, and application. *J Natl Cancer Inst*. 2005; 97:715–723. [PMID: 15900041]. [PubMed: 15900041]
17. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996; 15:361–387. [PMID: 8668867]. [PubMed: 8668867]
18. Vickers AJ, Cronin AM, Elkin EB, Gonen M. Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers. *BMC Med Inform Decis Mak*. 2008; 8:53. [PMID: 19036144]. [PubMed: 19036144]
19. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making*. 2006; 26:565–574. [PMID: 17099194]. [PubMed: 17099194]

20. Baker SG. Putting risk prediction in perspective: relative utility curves. *J Natl Cancer Inst.* 2009; 101:1538–1542. [PMID: 19843888]. [PubMed: 19843888]
21. Baker SG, Van Calster B, Steyerberg EW. Evaluating a new marker for risk prediction using the test tradeoff: an update. *Int J Biostat.* 2012;8. [PMID: 22499728].
22. Field JK, Liloglou T, Niaz A, Bryan J, Gosney JR, Giles T, et al. EUELC Collaborators. EUELC project: a multi-centre, multipurpose study to investigate early stage NSCLC, and to establish a biobank for ongoing collaboration. *Eur Respir J.* 2009; 34:1477–1486. [PMID: 19948914]. [PubMed: 19948914]
23. Garcia-Closas M, Kelsey KT, Wiencke JK, Xu X, Wain JC, Christiani DC. A case-control study of cytochrome P450 1A1, glutathione S-transferase M1, cigarette smoking and lung cancer susceptibility (Massachusetts, United States). *Cancer Causes Control.* 1997; 8:544–553. [PMID: 9242469]. [PubMed: 9242469]
24. Field JK, Smith DL, Duffy S, Cassidy A. The Liverpool Lung Project research protocol. *Int J Oncol.* 2005; 27:1633–1645. [PMID: 16273220]. [PubMed: 16273220]
25. Carel R, Olsson AC, Zaridze D, Szeszenia-Dabrowska N, Rudnai P, Lissowska J, et al. Occupational exposure to asbestos and man-made vitreous fibres and risk of lung cancer: a multicentre case-control study in Europe. *Occup Environ Med.* 2007; 64:502–508. [PMID: 17053017]. [PubMed: 17053017]
26. Reilly BM, Evans AT. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. *Ann Intern Med.* 2006; 144:201–209. [PMID: 16461965]. [PubMed: 16461965]
27. Gail MH, Pfeiffer RM. On criteria for evaluating models of absolute risk. *Biostatistics.* 2005; 6:227–239. [PMID: 15772102]. [PubMed: 15772102]
28. Kerlikowske K. Evidence-based breast cancer prevention: the importance of individual risk [Editorial]. *Ann Intern Med.* 2009; 151:750–752. [PMID: 19920276]. [PubMed: 19920276]
29. Barlow WE, White E, Ballard-Barbash R, Vacek PM, Titus-Ernstoff L, Carney PA, et al. Prospective breast cancer risk prediction model for women undergoing screening mammography. *J Natl Cancer Inst.* 2006; 98:1204–1214. [PMID: 16954473]. [PubMed: 16954473]
30. Rockhill B, Spiegelman D, Byrne C, Hunter DJ, Colditz GA. Validation of the Gail et al. model of breast cancer risk prediction and implications for chemo-prevention. *J Natl Cancer Inst.* 2001; 93:358–366. [PMID: 11238697]. [PubMed: 11238697]
31. Gail MH. Discriminatory accuracy from single-nucleotide polymorphisms in models to predict breast cancer risk. *J Natl Cancer Inst.* 2008; 100:1037–1041. [PMID: 18612136]. [PubMed: 18612136]
32. Apicella C, Dowty JG, Dite GS, Jenkins MA, Senie RT, Daly MB, et al. Validation study of the LAMBDA model for predicting the BRCA1 or BRCA2 mutation carrier status of North American Ashkenazi Jewish women. *Clin Genet.* 2007; 72:87–97. [PMID: 17661812]. [PubMed: 17661812]
33. Simmons RK, Sharp S, Boekholdt SM, Sargeant LA, Khaw KT, Wareham NJ, et al. Evaluation of the Framingham risk score in the European Prospective Investigation of Cancer-Norfolk cohort: does adding glycated hemoglobin improve the prediction of coronary heart disease events? *Arch Intern Med.* 2008; 168:1209–1216. [PMID: 18541829]. [PubMed: 18541829]
34. Brindle P, Beswick A, Fahey T, Ebrahim S. Accuracy and impact of risk assessment in the primary prevention of cardiovascular disease: a systematic review. *Heart.* 2006; 92:1752–1759. [PMID: 16621883]. [PubMed: 16621883]
35. Brindle PM, McConnachie A, Upton MN, Hart CL, Davey Smith G, Watt GC. The accuracy of the Framingham risk-score in different socioeconomic groups: a prospective study. *Br J Gen Pract.* 2005; 55:838–845. [PMID: 16281999]. [PubMed: 16281999]
36. Vergouwe Y, Moons KG, Steyerberg EW. External validity of risk models: Use of benchmark values to disentangle a case-mix effect from incorrect coefficients. *Am J Epidemiol.* 2010; 172:971–980. [PMID: 20807737]. [PubMed: 20807737]
37. Steyerberg, EW. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating.* New York: Springer; 2009.

38. van Klaveren RJ, de Koning HJ, Mulshine J, Hirsch FR. Lung cancer screening by spiral CT. What is the optimal target population for screening trials? *Lung Cancer*. 2002; 38:243–252. [PMID: 12445745]. [PubMed: 12445745]
39. Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults. Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III). *JAMA*. 2001; 285:2486–2497. [PMID: 11368702]. [PubMed: 11368702]
40. Pepe MS, Janes HE. Gauging the performance of SNPs, biomarkers, and clinical factors for predicting risk of breast cancer [Editorial]. *J Natl Cancer Inst*. 2008; 100:978–979. [PMID: 18612128]. [PubMed: 18612128]
41. Baldwin DR, Duffy SW, Wald NJ, Page R, Hansell DM, Field JK. UK Lung Screen (UKLS) nodule management protocol: modelling of a single screen randomised controlled trial of low-dose CT screening for lung cancer. *Thorax*. 2011; 66:308–313. [PMID: 21317179]. [PubMed: 21317179]
42. Hung RJ, McKay JD, Gaborieau V, Boffetta P, Hashibe M, Zaridze D, et al. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature*. 2008; 452:633–637. [PMID: 18385738]. [PubMed: 18385738]
43. Truong T, Hung RJ, Amos CI, Wu X, Bickebo`ller H, Rosenberger A, et al. Replication of lung cancer susceptibility loci at chromosomes 15q25, 5p15, and 6p21: a pooled analysis from the International Lung Cancer Consortium. *J Natl Cancer Inst*. 2010; 102:959–971. [PMID: 20548021]. [PubMed: 20548021]

Appendix: The EUELC Consortium

Christian Brambilla, Institut Albert Bonniot, Université Joseph Fourier, INSERM U823, Grenoble, France

Yves Martinet, Centre Hospitalier Universitaire de Nancy, Nancy, France

Frederik B. Thunnissen, Department of Pathology, Canisius Wilhelmina Ziekenhuis, Nijmegen, and Department of Pathology, Vrije Universiteit Medical Center, Amsterdam, the Netherlands

Peter J. Snijders, Department of Pathology, Vrije Universiteit Medical Center, Amsterdam, the Netherlands

Gabriella Sozzi, Department of Experimental Oncology and Laboratories, Fondazione Istituto Di Ricovero e Cura a Carattere Scientifico Istituto Nazionale Tumori, Milan, Italy

Angela Risch, German Cancer Research Centre, Heidelberg, Germany

Heinrich D. Becker, Thoraxklinik at Heidelberg University, Heidelberg, Germany

J. Stuart Elborn, Respiratory Medicine Research Group, Centre for Infection and Immunity, Queen's University, Belfast, United Kingdom

Luis M. Montuenga, Center for Applied Medical Research, University of Navarra, Navarra, Spain

Ken J. O'Byrne, St. James Hospital, Dublin, Ireland

David J. Harrison, University of Edinburgh, Edinburgh, United Kingdom

Jacek Niklinski, Medical Academy of Bialystok, Bialystok, Poland

John K. Field, Roy Castle Lung Cancer Research Programme, The University of Liverpool
Cancer Research Centre, Institute of Translational Medicine, The University of Liverpool,
Liverpool, United Kingdom

Appendix Table 1

Revised LLP Risk Model Regression Coefficients and Internal Discriminative Validation *

Variable	Missing Risk Factor				
	Smoking Duration	History of Pneumonia	History of Cancer	Family History of Lung Cancer	Asbestos Exposure
Model risk factors					
Smoking duration					
1–19 y	–	0.7908	0.7526	0.8159	0.7950
20–39 y	–	1.4906	1.4562	1.4743	1.4704
40–59 y	–	2.5216	2.5274	2.5708	2.5359
60 y	–	2.6762	2.7624	2.7924	2.7424
History of pneumonia	0.5619	–	0.5893	0.5814	0.5917
History of cancer	0.8599	0.6620	–	0.6808	0.6795
Family history of lung cancer					
Early onset (age <60 y)	0.9351	0.6538	0.7096	–	0.6901
Late onset (age ≥ 60 y)	0.2779	0.1850	0.1943	–	0.2009
Asbestos exposure	0.6659	0.6273	0.6367	0.6323	–
Internal validation					
AUC	0.6374	0.7554	0.7543	0.7537	0.7526

AUC — area under the receiver-operating characteristic curve; LLP = Liverpool Lung Project.

* Estimated from the LLP case-control study.

Appendix Table 2

Projected 5-Year Absolute Risk for a Person Aged 66 Years With Different Risk Factor Profiles, Using the LLP Risk Model *

Smoking Duration by Sex	Family History of Lung Cancer [†]	History of Cancer	History of Pneumonia	Asbestos Exposure	5-Year Absolute Risk%
Male					
45 y	Early onset	Yes	Yes	Yes	41.04
0 y	Early onset	Yes	Yes	Yes	5.37
45 y	None	No	No	No	4.84
0 y	None	No	No	No	0.41

Smoking Duration by Sex	Family History of Lung Cancer [†]	History of Cancer	History of Pneumonia	Asbestos Exposure	5-Year Absolute Risk%
45 y	Early onset	No	No	No	9.33
45 y	None	Yes	No	No	9.09
45 y	None	No	Yes	No	8.51
45 y	None	No	No	Yes	8.76
Female					
45 y	Early onset	Yes	Yes	Yes	32.8
0 y	Early onset	Yes	Yes	Yes	3.83
45 y	None	No	No	No	3.45
0 y	None	No	No	No	0.21
45 y	Early onset	No	No	No	6.73
45 y	None	Yes	No	No	6.55
45 y	None	No	Yes	No	6.12
45 y	None	No	No	Yes	6.31

* Table adapted from reference 12. This table shows that risk factors other than smoking can confer a substantial risk for both sexes, which demonstrates the importance of such factors in risk prediction. The diversity of predicted risks for a smoker and nonsmoker of similar age with or without other risk factors can be appreciated. For example, a man aged 66 y with no smoking history but a family history of early-onset cancer and a history of pneumonia, other cancer, and asbestos exposure has a 5-y risk of 5.37%. This is higher than the 5-y risk of 4.84% for a man of similar age who has smoked for 45 y and has no other risk factors.

[†] Early onset is age <60 y; late onset is age ≥ 60 y

Appendix Figure 1. Comparison of observed and predicted cases of cancer, by quartile of predicted risk in the LLPC data.

Improved model calibrations were observed at higher quartiles of predicted absolute risks, which demonstrates a better calibration of the model at risk probabilities at which screening might be initiated. LLPC = Liverpool Lung Project prospective cohort.

Appendix Table 3

Patient Risk Classification, by Status

Absolute Risk Threshold	Risk Group	EUELC (n= 1868)		Harvard (n= 2922)		LLPC (n= 7652)	
		Case Patients (n= 585) n (%)	Control Participants (n= 1283) n (%)	Case Patients (n= 1738) n (%)	Control Participants (n= 1184) n (%)	With Disease (n= 420) n (%)	Disease-Free (n= 7232) n (%)
2.5%	<2.5%	262 (44.8)	895 (69.8)	613 (35.3)	865 (73.1)	108 (25.7)	4873 (67.4)
	2.5%	323 (55.2)	388 (30.2)	1125 (64.7)	319 (26.9)	312 (74.3)	2359 (32.6)
5.0%	<5.0%	376 (64.3)	1082 (84.3)	880 (50.6)	1031 (87.1)	179 (42.6)	5867 (81.1)
	5.0%	209 (35.7)	201 (15.7)	858 (49.4)	153 (12.9)	241 (57.4)	1365 (18.9)
10.0%	<10.0%	502 (85.8)	1210 (94.3)	1299 (74.7)	1133 (95.7)	303 (72.1)	6690 (92.5)
	10.0%	83 (14.2)	73 (5.7)	439 (25.3)	51 (4.3)	117 (27.9)	542 (7.5)

EUELC — European Early Lung Cancer; LLPC = Liverpool Lung Project prospective cohort.

Appendix Figure 2. Relative utility of the LLP risk model and smoking duration in the EUELC and Harvard data sets.

At high risk thresholds, the LLP model showed moderate relative utility compared with smoking duration in the EUELC and Harvard data sets. The good performance at higher predicted absolute risks (greater than the probability of disease) are relevant for a default strategy of screening no one. The smoking duration for relative utility curves involves 4 categories. EUELC = European Early Lung Cancer; LLP = Liverpool Lung Project.

Context

A limitation in the development of models to predict lung cancer risk has been a lack of focus on predicted clinical benefit. Models should balance accuracy and the potential harms and benefits of screening for disease at different risk thresholds.

Contribution

In this decision analysis of data from 3 study populations, the Liverpool Lung Project risk model performed well in predicting clinical benefit across a range of thresholds of possible risk compared with other approaches to risk stratification.

Implication

Such analyses may help in planning the implementation of lung cancer screening programs.

—*The Editors*

Highlights of www.annals.org

Online-first articles: The latest clinical news before it is published in the print edition.

Quick links: Article-specific CME, slides, patient information, multimedia, and commenting and sharing features.

Smarter article collections: Immediate access to collections organized by specialty, disease, and special topics.

Better search tools: More meaningful results, with screening and refining options.

Mobility: Automatic mobile-friendly display when accessing annals.org from any smart phone.

Figure 1. Distribution of predicted absolute risk for participants in the EUELC, Harvard, and LLPC studies

EUELC = European Early Lung Cancer; LLPC = Liverpool Lung Project prospective cohort.

Figure 2. Decision curves and relative utility curve of the LLP risk model compared with alternative strategies for lung cancer screening decisions

The smoking duration for relative utility curves involves 4 categories, whereas it had to be restricted to 2 categories for decision curves for technical reasons. DCA = decision curve analysis; EUELC = European Early Lung Cancer; LLP = Liverpool Lung Project; LLPC = Liverpool Lung Project prospective cohort; RU = relative utility.

Table 1

Risk Factor Distributions and Estimated Model Coefficients From the LLP Case-Control Data*

Variable	Case Patients, <i>n</i> (%)	Control Participants, <i>n</i> (%)	Odds Ratio (95% CI)	Full Model	Reduced Model [†]
Risk factor					
Smoking duration					
Never	25 (5.9)	240 (28.3)	1.00 (reference)	–	–
1–19 y	37 (8.7)	171 (20.2)	2.17 (1.21–3.85)	0.769	0.795
20–39 y	122 (28.6)	249 (29.4)	4.27 (2.62–6.94)	1.452	1.470
40–59 y	227 (53.2)	176 (20.8)	12.27 (7.41–20.30)	2.507	2.536
60 y	16 (3.8)	12 (1.4)	15.25 (5.71–40.65)	2.724	2.742
History of pneumonia	97 (22.7)	129 (15.2)	1.83 (1.26–2.64)	0.602	0.592
Asbestos exposure	150 (35.1)	202 (23.8)	1.89 (1.35–2.62)	0.634	–
History of cancer	60 (14.1)	59 (7.0)	1.96 (1.22–3.14)	0.675	0.680
Family history of lung cancer					
None	314 (73.5)	699 (82.4)	1.00 (reference)	–	–
Early onset (age <60 y)	44 (10.3)	40 (4.7)	2.02 (1.18–3.45)	0.703	0.690
Late onset (age ≥60 y)	69 (16.2)	109 (12.9)	1.18 (0.79–1.76)	0.168	0.201
Internal validation					
AUC	–	–	–	0.758	0.753

AUC = area under the receiver-operating characteristic curve; LLP = Liverpool Lung Project.

* Table adapted from reference 12.

[†] Excludes asbestos exposure.

Table 2
Epidemiologic and Clinical Characteristics of Patients in the 3 Validation Data Sets

Risk Factor	Case-Control Studies				LLPC	
	EUELC (n= 1868)		Harvard (n= 2922)		Cohort Size, n (%)	Cohort With Lung Cancer Incidence, n (%)
	Case Patients, n (%)	Control Participants, n (%)	Case Patients, n (%)	Control Participants, n (%)		
All participants	585 (31.3)	1283 (68.7)	1738 (59.5)	1184 (40.5)	7652	420 5.5
Sex						
Male	448 (76.6)	956 (74.5)	945 (54.4)	546 (46.1)	3587 (46.9)	234 (55.7) 6.5
Female	137 (23.4)	327 (25.5)	793 (45.6)	638 (53.9)	4065 (53.1)	186 (44.3) 4.6
Age						
<50 y	35 (5.9)	97 (7.6)	137 (7.9)	215 (18.2)	803 (10.5)	10 (2.4) 1.3
50–59 y	178 (30.4)	401 (31.3)	370 (21.3)	289 (24.4)	2481 (32.4)	71 (16.9) 2.9
60–69 y	221 (37.8)	491 (38.3)	621 (35.7)	413 (34.9)	2890 (37.8)	172 (41.0) 6.0
70–79 y	139 (23.8)	265 (20.7)	520 (29.9)	245 (20.7)	1451 (18.9)	152 (36.2) 10.5
80 y	12 (2.1)	29 (2.3)	90 (5.2)	22 (1.9)	27 (0.4)	15 (3.6) 55.6
Smoking duration						
Never	41 (7.0)	361 (28.1)	114 (6.6)	396 (33.5)	2783 (36.4)	31 (7.4) 1.1
1–19 y	33 (5.6)	181 (14.1)	122 (7.0)	207 (17.5)	1302 (17.0)	29 (6.9) 2.2
20–39 y	226 (38.6)	418 (32.6)	646 (37.2)	391 (33.0)	2013 (26.3)	122 (29.1) 6.1
40–59 y	270 (46.2)	307 (23.9)	797 (45.9)	182 (15.4)	1496 (19.6)	223 (53.1) 14.9
60 y	15 (2.6)	16 (1.3)	59 (3.4)	8 (0.7)	58 (0.8)	15 (3.6) 25.9
History of pneumonia						
No	486 (83.1)	1112 (86.7)	931 (53.6)	888 (75.0)	6707 (87.7)	360 (85.7) 5.4

Risk Factor	Case-Control Studies				LLPC		
	EUELC (n= 1868)		Harvard (n= 2922)		Cohort Size, n (%)	Lung Cancer Incidence, n (%)	Cohort With Disease, %
	Case Patients, n (%)	Control Participants, n (%)	Case Patients, n (%)	Control Participants, n (%)			
Yes	99 (16.9)	171 (13.3)	807 (46.4)	296 (25.0)	945 (12.3)	60 (14.3)	6.4
History of cancer							
No	463 (79.2)	1144 (89.2)	947 (54.5)	736 (62.2)	6802 (88.9)	297 (70.7)	4.4
Yes	122 (20.9)	139 (10.8)	791 (45.5)	448 (37.8)	852 (11.1)	123 (29.3)	14.5
Asbestos exposure*							
No	468 (80.0)	1209 (94.2)	1335 (76.8)	1171 (98.9)	–	–	–
Yes	117 (20.0)	74 (5.8)	403 (23.2)	13 (1.1)	–	–	–
Family history of lung cancer							
None	480 (82.1)	1115 (86.9)	1446 (83.2)	1045 (88.3)	6161 (80.5)	352 (83.8)	5.7
Early onset (age <60 y)	69 (11.8)	101 (7.9)	174 (10.1)	74 (6.3)	973 (12.7)	39 (9.3)	4.0
Late onset (age ≥ 60 y)	36 (6.2)	67 (5.2)	118 (6.8)	65 (5.5)	518 (6.7)	29 (6.9)	5.6

EUELC = European Early Lung Cancer; LLPC = Liverpool Lung Project prospective cohort.

*Data on asbestos exposure were missing for 7387 LLPC participants (96.5%).

Table 3

Sensitivity, Specificity, Net Benefit, Net Gain, and AUC of the LLP Risk Model for Specified Risk Thresholds in Each of 3 Populations

Variable and Measure	Study Population		
	EUELC	Harvard	LLPC
Absolute risk threshold*			
2.5%			
Sensitivity, %	55.2	64.7	74.3
Specificity, %	69.8	73.1	67.4
Accuracy, %	65.2	68.1	67.8
Net benefit	0.0190	0.0345	0.0261
Net gain, % [†]	1.1	0.9	0.53
5.0%			
Sensitivity, %	35.7	49.4	57.4
Specificity, %	84.3	87.1	81.1
Accuracy, %	69.1	64.7	79.8
Net benefit	0.0127	0.0249	0.0171
Net gain, % [†]	3.0	2.52	2.3
10.0%			
Sensitivity, %	14.2	25.3	27.9
Specificity, %	94.3	95.7	92.5
Accuracy, %	69.2	53.8	89.0
Net benefit	0.0055	0.0134	0.0060
Net gain, % [†]	7.65	6.87	6.78
Discriminative performance			
AUC (95% CI)	0.67 (0.64–0.69)	0.76 (0.75–0.78)	0.82 (0.80–0.85)

AUC = area under the receiver-operating characteristic curve; EUELC = European Early Lung Cancer; = Liverpool Lung Project; = Liverpool Lung Project prospective cohort.

* Absolute risk for lung cancer at which screening is initiated.

[†] Gain in net benefit for LLP risk model over a screen-all strategy.