Evaluation of the accuracy of the PLCO(m2012) 6-year lung cancer risk prediction model in the CARTaGENE population-based cohort

Rodolphe Jantzen, MD[1,2,*], Nicole Ezer MD MPH[3], Sophie Camilleri-Broët MD PhD[3], Martin C. Tammemägi PhD[4], Philippe Broët, MD PhD[1,2,5,6,*]

[1] CARTaGENE, Research Center, CHU Sainte-Justine, Montreal, Quebec, Canada

[2] Université de Montréal, Montréal, Québec, Canada

[3] McGill University, Montréal, Canada

[4] Prevention and Cancer Control, Ontario Health (Cancer Care Ontario), Toronto, Ontario, Canada; Department of Health Sciences, Brock University, St. Catharines, Ontario, Canada.

[5] University Paris-Saclay, CESP, INSERM, Villejuif, France

[6] Assistance Publique-Hôpitaux de Paris (AP-HP), Hôpitaux Universitaires Paris-Sud, Hôpital Paul Brousse, 12 Avenue Paul Vaillant Couturier, 94807 Villejuif, France

*** Correspondence:**

Keywords: lung cancer, validation study, PLCOm2012, screening

Abstract

Background: The first objective was to validate PLCOm2012 lung cancer risk prediction tool in Quebec. The second objective was to determine the hypothetical performances of seven screening strategies (2021 United-States and 2016 Canadian screening recommendations, PLCOm2012 with ≥1.51%/6y, ≥1.7%/6y and ≥2.0%/6y thresholds, Quebec pilot criteria with 55-74 and 50-74 years ranges).

Methods: The population-based CARTaGENE cohort was used, excluding participants who never smoked, had missing smoking status or had a lung cancer history. Lung cancer risk was calculated at inclusion for 11,652 participants and prediction calibration and discrimination were assessed. For

investigating screening strategies, 8,938 participants were included. Strategies were evaluated between 1998 and 2015 using shift and serial scenarios (PLCOm2012 eligibility estimated annually and each 6 years, respectively).

Results: Objective 1. 176 (1.51%) lung cancers diagnosed. Expected-to-observed ratio was 0.68 [95%CI 0.59-0.79] and C-statistic was 0.727 [0.679-0.770]. For the selected risk thresholds, sensitivities ranged from 44.9% [37.4-52.6] to 52.3% [44.6-59.8], specificities from 81.6% [80.8-82.3] to 87.7% [87-88.3] and positive predictive values from 4.2% [3.4-5.1] to 5.3% [4.2-6.5]. Objective 2. Using shift scenario, cancers detected with Quebec pilot criteria were lower than PLCOm2012 at ≥2.0%/6y threshold (48.3% vs 50.2%) for a similar number of scans per detected cancer. Serial scenarios led to less lung cancers detected (up to 26 less) but higher positive predictive values, the highest being PLCOm2012 at 2% (6.0% [4.8-7.3]).

Interpretation: PLCOm2012 had good discrimination but weak calibration in Quebec. It may be helpful to adjust the PLCOm2012 intercept to improve calibration in the Quebec setting.

1 Introduction

Lung cancer remains the leading cause of cancer-related mortality in Canada and worldwide (18%) (1). Two large scaled randomized controlled trials of lung cancer screening, the National Lung Screening Trial (NLST) and NELSON trial (2,3), have conclusively demonstrated efficacy, with a reduction of lung cancer mortality in men of 8-26%, and 26-41% in women in high-risk smokers. This has been demonstrated to be cost-effective (4). As a consequence, screening is now widely recommended, but implementation remains limited (5), and varies across countries.

Multiple predictive models for lung cancer have been developed with different predicted outcomes (incidence, death), prediction horizon (1 year, 6 years) and risk factors included (5). Among them, the Tammemägi *et al.* lung cancer predictive model (PLCOm2012) (6) showed good discriminative power (area under the ROC curve (AUC) around 0.8). It has been externally validated across different countries (5–8), and most recently in the International Lung Screen Trial (Australia, Canada, Hong Kong, the UK. and Spain) to prospectively identify the best screening strategy between national guidelines and risk prediction model (9). Provisional findings (10) showed that PLCOm2012 detected significantly more lung cancers than the 2013 United States Preventive Services Task Force (USPSTF) criteria. Moreover, PLCOm2012 was better than the 2013 USPSTF for the sensitivity, death averted, screening efficiency and reducing race and sex disparities (11–13).

In Canada, the Canadian Task Force on Preventive Health Care (CTFPHC) (14) recommend screening for lung cancer using the NLST entry criteria (55 to 74 years, at least a 30 pack-year smoking history, smoking quit-time less than 15 years) with low-dose computed tomography scan (LDCT) every year for three consecutive screens. Quebec and Ontario have used PLCOm2012 with a 2% threshold and age of 55-74 years, based on cost-effectiveness analyses (15,16), to select residents for screening. In other provinces, screening is in the process of being implemented.

Risk models may be prone to increase the selection of older individuals with more comorbidities and competing causes of death. The role of competing causes of death in these high-risk smokers, and baseline risk factors in the population, may affect the performance of these models in different

countries and may vary across provinces. Binary criteria can lead to the selection of individuals with too low risks to benefit from screening (17).

Here, we validate the PLCOm2012 model at six years in the CARTaGENE Quebec population-based cohort. We also report the efficiency of seven screening strategies that differed in criteria, frequency of risk score calculation (each year or 6 years) and risk score thresholds, if theoretically applied between 1998 and 2015 to our Quebec cohort.

2 Methods

2.1 Population studied and definition of lung cancer

This study used the population-based cohort CARTaGENE recruited in phase A (2009-2010), composed of 19,985 Quebec residents aged between 40 and 69 years described previously (18). Briefly, participants were randomly selected to be broadly representative of the population in metropolitan areas. Questionnaires at enrollment included data on age, ethnicity, education, body mass index, self-reported history of chronic obstructive pulmonary disease (COPD), familial history of lung cancer, smoking status, cigarettes per day at inclusion and when the participant smoked the most, start and stop smoking years, smoking duration, and duration of smoking cessation. Participants can be linked with the Quebec administrative health databases from 1998-2015, providing data on cancer diagnoses. Individuals who never smoked or had missing smoking data or a lung cancer diagnosed before 1998 were excluded (Figure 1).

The Tonelli *et al.* algorithm was used to define an incident lung cancer case (19) (i.e., individuals with at least 2 claims in 2 years or one hospitalization related to a lung cancer; incidence date was the date of first hospital discharge or first claim).

2.2 Study design

Our first objective was to externally validate the PLCOm2012 model for estimating the 6-year absolute risk of lung cancer from enrollment in the CARTaGENE cohort. The second objective was to determine the hypothetical performances of 7 different screening strategies to detect lung cancers if applied between January 1st 1998 and December 31st 2015: the original PLCOm2012 model (6) using ≥1.51%/6y, ≥1.7%/6y and ≥2.0%/6y thresholds (11); the 2021 USPSTF criteria (20); the 2016 CTFPHC criteria (14); the Quebec pilot criteria (PLCOm≥2% and 55-74 years range) (21) and the Quebec pilot with a 50-74 years range to test the lower age threshold of the 2021 USPSTF criteria.

2.3 Statistical analysis

For both objectives, education, family history, smoking status and COPD status were considered unchanged after enrollment in the cohort. Missing data for variables in the PLCOm2012 model were replaced by the centering value of the PLCOm2012 original article for continuous variables (Table 2 in (6)) and the mode for categorical variables. The proportion of missing data was higher for smoking-related variables such as intensity and duration, but was limited (Supplementary Table S1).

2.3.1 Six-year risk prediction accuracy for lung cancer from enrollment in the CaG cohort

The individual 6-year absolute risk of lung cancer was estimated using the coefficients and variables according to the Tammemägi *et al.* original article (6) (Supplementary Materials). Participants with a lung cancer occurrence before the inclusion date were excluded (Figure 1).

We computed the expected-to-observed ratio (E/O) with 95% confidence intervals (95%CI), from the sum of the estimated risk (i.e., the number of expected cases) divided by the number of observed cases. We determined the E/O with 95%CI in four absolute risk groups: < 0.5%, 0.5%-1%, 1-2%, and >2%. The best calibrated models have an estimate closer to 1. Calibration graphs were plotted to compare in each risk group the proportion of observed cases of lung cancer at 6-years using a Kaplan-Meier estimator, and the proportion of expected cases (i.e., median risk). Global discrimination was assessed by the C-statistic with an inverse probability of censoring weighting estimation of cumulative time-dependent ROC curve with their 95%CI (22–24). We calculated sensitivity, specificity and positive predictive value for ≥1.51%/6y, ≥1.7%/6y and ≥2.0%/6y thresholds. We also plotted predictiveness curves (i.e., the risk quantile against the corresponding cumulative proportion of the population with risks below this quantile).

2.3.2 Hypothetical efficiency of 7 lung cancer screening strategies

For this objective, the efficiency of a lung cancer screening strategy was assessed as if it has been implemented between January 1st 1998 and December 31st 2015. Participants with missing cigarettes per day, missing start/stop smoking date, or with a stop smoking date occurring before the start smoking date were excluded (Figure 1). We considered a lung cancer to be "screen detected" if the participant was eligible for screening and if a LDCT would have been theoretically performed in the year before the actual cancer occurrence date. To have at least one-year post-screening for each participant, the last occurrence of what we considered a "screening" was made in 2014.

A participant was "eligible for screening" if he/she met eligibility criteria of the considered screening strategy. For the binary screening scenarios (USPSTF and CTFPHC), we determined eligibility yearly. For the screening scenarios based on the PLCOm2012 model, we determined eligibility based on two schemes: 1) the *shift scenario* where we estimated eligibility *annually* using the PLCOm2012 thresholds (and added age for the Quebec program); 2) *the serial scenario* where we determined eligibility *at 6-year intervals*, starting in 1998 for the models using PLCO risk criteria, and additionally, starting when the participant was 50 years of age or 55 years of age for the Quebec pilot strategies (Figure 2).

For each of the seven screening strategies, we calculated the total number of participants theoretically eligible to be screened, the total number of LDCT that would have been performed, the number of incident lung cancers that would have been detected, the number of LDCT to be performed to detect one lung cancer, the number of participants to be screened to detect one lung cancer. We also estimated the number of LDCT per participant that would have been performed before the detection of the lung cancer, and the number of LDCT per cancer-free participants with at least one LDCT.

We calculated the sensitivity, specificity and positive predictive value with their 95%CI. The sensitivity was the probability of being screened in the year prior to a lung cancer being diagnosed. For the specificity, we reported the probability of being not eligible per cancer-free participant; for the positive predictive value, we reported the probability to detect a lung cancer for a participant being screened and having at least one LDCT.

All statistical analyses were performed using R software, version 4.0 (25).

3 Results

3.1 Six-year risk prediction accuracy for lung cancer from enrollment in the CaG cohort

The 11,652 participants included in the cohort used for external validation of the PLCOm2012 model had a median age at inclusion of 54 years [IQR 49-61] and a median follow-up time of 5.9 years [IQR 5.7-6]. There were 176 (1.51%) lung cancers diagnosed during the 6-year period following enrollment. For the PLCOm2012 model, 19.0%, 16.2% and 12.8% of the cohort had a 6-year lung cancer risk estimated to equal or higher than 1.51%, 1.7% and 2%, respectively (Figure 3A). The estimated 6-year lung cancer median risk scores were 1.67% [IQR 0.62-3.86] and 0.54% [IQR 0.27-1.16] for the participants with and without a diagnosis of lung cancer, respectively.

The global calibration was 0.68 [95%CI 0.59-0.79]. All E/O were lower than one in each risk group, the significant ones being in the <0.5% and ≥2% groups (0.37 [0.27-0.51] and 0.74 [0.59-0.92], respectively) (Figure 3B). The c-statistic was 0.727 [0.679-0.770] (Figure 3C). For the different thresholds, the sensitivity was ranging from 44.9% [37.4-52.6] to 52.3% [44.6-59.8]. The specificity was ranging from 81.6% [80.8-82.3] to 87.7% [87-88.3]. The positive predictive value was ranging from 4.2% [3.4-5.1] to 5.3% [4.2-6.5] (Table 1).

3.2 Hypothetical efficiency of 7 lung cancer screening strategies

Among the 8,938 participants included to compare the efficiency of the 7 lung cancer screening strategies, 205 (2.3%) had a lung cancer between 1998 and 2015.

3.2.1 Shift scenario

The number of LDCTs performed ranged from 15,201 (Quebec pilot) to 40,448 (USPSTF), while the number of cancers detected ranged from 99 (48.3%) (Quebec pilot) to 133 (64.9%) (USPSTF) (Table 2). The PLCOm2% detected more lung cancers than CTFPHC with fewer scans. The number of lung cancers detected using the Quebec pilot criteria (greater than 2% risk and age) were lower than those where the PLCOm2012 risk threshold of ≥2%/6y alone (99 (48.3%) vs 103 (50.2%)) for a similar number of CTs performed to detect one lung cancer (153.5 vs 162.9). The number of screened participants needed to detect one cancer was the lowest for the Quebec pilot (19.5) and highest for the USPSTF (33.4). The USPSTF had the highest sensitivity (64.9% [57.9-71.4]), while the Quebec pilot had the highest positive predictive value per screened (5.1% [4.2-6.2]). The results for the Quebec pilot criteria with an 50-74 age range were similar to a strategy that use only PLCOm2012 2% eligibility cut-off alone (Table 2).

Using the CTFPHC and USPSTF strategies, 11 and 13 participants stopped their screening before the detection of their lung cancer, respectively, as they stopped smoking for more than 15 years. Their lung cancer occurred between 3.8 and 14.5 years after the last LDCT. Among these 13 participants, 4 were detected by the PLCOm2012 models. None of the participants stopped his/her screening before their lung cancer using the other strategies.

3.2.2 Serial scenario

Among the lung cancers that occurred between 1998 and 2015, the number of detected cancers using the serial scenario was lower compared to the shift scenario, ranging from 16 (PLCOm1.51%) to 26 (Quebec pilot 50-74 years) less cancers detected. The number of screened participants needed to detect one cancer was close for the Quebec pilot and lower for PLCOm2012. The sensitivities were all lower, while the positive predictive values were higher, the highest being the PLCOm2% (6.0% [4.8-7.3]) (Table S2 in Supplementary Appendix).

4 Interpretation

We validated the PLCOm2012 model to predict lung cancer at 6 years in the CARTaGENE cohort. We also assessed different lung cancer screening strategies in a theoretical screening program between 1998 and 2015.

4.1 Six-year risk prediction accuracy for lung cancer from enrollment in the CaG cohort

This is the first time that the PLCOm2012 model was validated in the Quebec population, even though a Quebec pilot study is assessing prospectively the PLCOm2012 model for lung cancer screening (21). In our cohort where the lung cancer incidence was 1.51%, the PLCOm2012 model underestimated the number of cases. This underestimation can be explained by the higher age-standardized lung cancer's incidence rate in Quebec (106.7 [95%CI 103.3-110.3] cases per 100,000 in 2010) than in the United-States (88.8 [95%CI 88.3-89.3]), based on information retrieved from national cancer databases (26–28). The risk was overestimated in the UK Biobank, EPIC-UK and Generation Study (incidence lower than 0.7%, E/O around 1.3) (8). In an Australian population-based cohort with excellent calibration, the lung cancer incidence rate was of 1.17% (7).

The AUCs of the PLCOm2012 were close to 0.80 in the PLCO cancer screening trial, in UK cohorts and in an Australian population-based cohort (6–8,29), which is higher than in our cohort (0.73). These differences might be explained by the statistical method used for assessing discrimination (c-statistic in our study), and how missing data were handled in UK and Australian cohorts (participants with missing data excluded, multiple imputation performed with more than 30% of missing data).

Compared to other studies, the Quebec population had lower sensitivity values when using the classical thresholds (7,11,29). Our positive predictive values were higher (7), which could be explained by the higher specificities and the higher prevalence of the lung cancer in our cohort. This may be due to smoking exposures in Quebec, which are known to be among the highest in North America.

4.2 Hypothetical efficiency of 7 lung cancer screening strategies

Re-assessing screening eligibility each 6 years instead of annually led to far less lung cancers being detected and lower sensitivities. However, the higher positive predictive values and the lower cost should be taken into account for public health policies. As the published studies seem to under-investigate the frequency with which lung cancer should be screened, it is necessary to evaluate different screening scenarios with non-binary predictive scores such as PLCOm2012.

The CTFPHC criteria did not perform well. Using PLCOm2012 with a 2% threshold with or without a 50-74 year age range detected more lung cancer per LDCT. The 2021 USPSTF criteria detected more cancers than the other strategies, but with far more LDCT screens performed and a significantly lower positive predictive value ($p<10^{-5}$). In the Pasquinelli *et al.* retrospective study (11), the 2013 USPSTF had a lower sensitivity than the 2021 USPSTF in our cohort (62.4% vs 64.9%). Moreover, CTFPHC and USPSTF criteria lead to some screening stopped before the occurrence of the lung cancer due to more than 15 years of smoking cessation. Therefore, the 2021 USPSTF seemed less efficient than predictive scores, but a more precise cost-effectiveness study is necessary.

The current Quebec pilot program, being deployed in 8 hospital centers in the province as of June 2021, is using the PLCOm2% with an age range from 55 to 74 years, preventing from screening older individuals that might not benefit from screening (e.g., more comorbidities and competing causes of death). Decreasing the age limit from 55 to 50 years, as suggested by the 2021 USPSTF recommendations, led to 4% absolute increase in the number of lung cancer detected. This was equivalent to PLCOm2% regarding the number of cancers detected, with less LDCT performed and participants screened by the PLCOm2%. However, this last result must be analyzed with participant older than 75 years. While the positive predictive value of the 50-74 age range was slightly lower than the 55-74 age range, decreasing the age threshold would allow the detection of lung cancers among younger individuals. In the Pasquinelli *et al.* retrospective study based on an actual lung cancer case series (11), PLCOm2012 had better sensitivities than in our cohort (ranging from 60.6% to 70.5%). Finally, the PLCOm2% with a serial scenario had the best positive predictive value, which is an important measure for public health policies in absence of cost-effectiveness analysis.

4.3 Strengths and limitations

This is the first study of lung cancer risk estimating and validation with Canadian participants outside of a clinical trial. The CARTaGENE cohort is representative of the Quebec urban population of middle-aged and older adults. Moreover, the linkage with a provincial administrative health database ensures all lung cancer cases are captured, improving the accuracy of our data.

Some limitations were present. Firstly, we did not have participants older than 75 years. Secondly, we did not know how lung cancers were detected (e.g., participant under surveillance for lung nodules). Therefore, the incidence date may depend on unobserved factors that may lead to biased estimate. Thirdly, some self-reported variables were only available before the inclusion date (COPD, familial history of lung cancer, smoking status). Fourthly, we had imputed missing data in the PLCOm2012 model, particularly regarding smoking related variables, but the proportion was lower than in other large cohort studies (7,8). Finally, we considered that lung cancers were

detected if an LDCT was made one year before the cancer's incidence, which was arbitrary but similar across the investigated scenarios.

4.4 Conclusion

The PLCOm2012 risk prediction model is currently used to select individuals for lung cancers in Canada. This model has a good discrimination but weak calibration for the Quebec population. A simple modification of the intercept in the prediction model may be proposed for improving the calibration in this population having a higher lung cancer incidence, but should be externally validated. Using the PLCOm2012 model with a 2% threshold, with an estimation of lung cancer risk and screening eligibility each six years, has added value in a large-scale provincial program as compared to other screening strategies. As predictive scores such as PLCOm2012 are intended to estimate a risk over a defined period, it is necessary to evaluate different screening scenarios. Additionally, lowering the age of onset of screening to 50 from 55 years may be considered but would require further cost-effectiveness analyses.

5 References

1. International Agency for Research on Cancer, World Health Organization. Cancer fact sheet: lung cancers [Internet]. 2020 [cited 2021 Apr 22]. Available from: https://gco.iarc.fr/today/data/factsheets/cancers/15-Lung-fact-sheet.pdf

2. The National Lung Screening Trial Research Team. Reduced Lung-Cancer Mortality with Low-Dose Computed Tomographic Screening. New England Journal of Medicine. 2011 Aug 4;365(5):395–409.

3. Zhao YR, Xie X, de Koning HJ, Mali WP, Vliegenthart R, Oudkerk M. NELSON lung cancer screening study. Cancer Imaging. 2011 Oct 3;11(1A):S79–84.

4. Goffin JR, Flanagan WM, Miller AB, Fitzgerald NR, Memon S, Wolfson MC, et al. Cost-effectiveness of Lung Cancer Screening in Canada. JAMA Oncol. 2015 Sep 1;1(6):807.

5. Sands J, Tammemägi MC, Couraud S, Baldwin DR, Borondy-Kitts A, Yankelevitz D, et al. Lung Screening Benefits and Challenges: A Review of The Data and Outline for Implementation. Journal of Thoracic Oncology. 2021 Jan 1;16(1):37–53.

6. Tammemägi MC, Katki HA, Hocking WG, Church TR, Caporaso N, Kvale PA, et al. Selection Criteria for Lung-Cancer Screening. N Engl J Med. 2013 Feb 21;368(8):728–36.

7. Weber M, Yap S, Goldsbury D, Manners D, Tammemagi M, Marshall H, et al. Identifying high risk individuals for targeted lung cancer screening: Independent validation of the PLCOm2012 risk prediction tool. Int J Cancer. 2017 Jul 15;141(2):242–53.

8. Robbins HA, Alcala K, Swerdlow AJ, Schoemaker MJ, Wareham N, Travis RC, et al. Comparative performance of lung cancer risk models to define lung screening eligibility in the United Kingdom. Br J Cancer. 2021 Apr 12;

9. Lim KP, Marshall H, Tammemägi M, Brims F, McWilliams A, Stone E, et al. Protocol and Rationale for the International Lung Screening Trial. Ann Am Thorac Soc. 2020 Apr;17(4):503–12.

10. Lam S, Myers R, Ruparel M, Atkar-Khattra S, Stone E, Manser R, et al. PL02.02 Lung Cancer Screenee Selection by USPSTF Versus PLCOm2012 Criteria – Interim ILST Findings. Journal of Thoracic Oncology. 2019 Oct;14(10):S4–5.

11. Pasquinelli MM, Tammemägi MC, Kovitz KL, Durham ML, Deliu Z, Rygalski K, et al. Risk Prediction Model Versus United States Preventive Services Task Force Lung Cancer Screening Eligibility Criteria: Reducing Race Disparities. Journal of Thoracic Oncology. 2020 Nov 1;15(11):1738–47.

12. ten Haaf K, Bastani M, Cao P, Jeon J, Toumazis I, Han SS, et al. A Comparative Modeling Analysis of Risk-Based Lung Cancer Screening Strategies. J Natl Cancer Inst. 2019 Nov 29;112(5):466–79.

13. Tammemägi MC, Church TR, Hocking WG, Silvestri GA, Kvale PA, Riley TL, et al. Evaluation of the Lung Cancer Risks at Which to Screen Ever- and Never-Smokers: Screening Rules Applied to the PLCO and NLST Cohorts. PLOS Medicine. 2014 Dec 2;11(12):e1001764.

14. Canadian Task Force on Preventive Health Care. Recommendations on screening for lung cancer. CMAJ. 2016 Apr 5;188(6):425–32.

15. Darling GE, Tammemägi MC, Schmidt H, Buchanan DN, Leung Y, McGarry C, et al. Organized Lung Cancer Screening Pilot: Informing a Province-wide Program in Ontario, Canada. The Annals of Thoracic Surgery [Internet]. 2020 Oct 7 [cited 2021 Mar 25];0(0). Available from: https://www.annalsthoracicsurgery.org/article/S0003-4975(20)31623-4/abstract

16. Tammemägi MC, Darling GE, Schmidt H, Llovet D, Buchanan DN, Leung Y, et al. Selection of individuals for lung cancer screening based on risk prediction model performance and economic factors – The Ontario experience. Lung Cancer. 2021 Jun 1;156:31–40.

17. Tammemägi MC. Selecting lung cancer screenees using risk prediction models—where do we go from here. Translational Lung Cancer Research [Internet]. 2018 Jun [cited 2021 Apr 29];7(3). Available from: https://tlcr.amegroups.com/article/view/21997

18. Awadalla P, Boileau C, Payette Y, Idaghdour Y, Goulet J-P, Knoppers B, et al. Cohort profile of the CARTaGENE study: Quebec's population-based biobank for public health and personalized genomics. International Journal of Epidemiology. 2013 Oct;42(5):1285–99.

19. Tonelli M, Wiebe N, Fortin M, Guthrie B, Hemmelgarn BR, James MT, et al. Methods for identifying 30 chronic conditions: application to administrative data. BMC Medical Informatics and Decision Making. 2015 Apr 17;15(1):31.

20. US Preventive Services Task Force, Krist AH, Davidson KW, Mangione CM, Barry MJ, Cabana M, et al. Screening for Lung Cancer: US Preventive Services Task Force Recommendation Statement. JAMA. 2021 Mar 9;325(10):962.

21. Canadian Partnership Against Cancer. Lung cancer screening in Canada: Environmental scan (2019-2020). 2020 Nov;41.

22. Uno H, Cai T, Tian L, Wei LJ. Evaluating Prediction Rules for t-Year Survivors with Censored Regression Models. Journal of the American Statistical Association. 2007;102(478):527–37.

23. Blanche P, Latouche A, Viallon V. Time-dependent AUC with right-censored data: a survey study. arXiv:12106805 [statME] [Internet]. 2012 Oct 25 [cited 2017 Sep 26]; Available from: http://arxiv.org/abs/1210.6805

24. Blanche P, Dartigues J-F, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. Statist Med. 2013 Dec 30;32(30):5381–97.

25. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2021. Available from: https://www.R-project.org/

26. Statistics Canada. Table 13-10-0111-01 Number and rates of new cases of primary cancer, by cancer type, age group and sex [Internet]. Government of Canada; [cited 2019 Nov 10]. Available from: doi.org/10.25318/1310011101-eng

27. Statistics Canada. Table 17-10-0005-01 Population estimates on July 1st, by age and sex [Internet]. Government of Canada; [cited 2019 Nov 10]. Available from: doi.org/10.25318/1710000501-eng

28. U.S. Cancer Statistics Working Group. U.S. Cancer Statistics Data Visualizations Tool, based on 2020 submission data (1999-2018): U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute [Internet]. 2021 [cited 2021 Sep 16]. Available from: www.cdc.gov/cancer/dataviz

29. ten Haaf K, Jeon J, Tammemägi MC, Han SS, Kong CY, Plevritis SK, et al. Risk prediction models for selection of lung cancer screening candidates: A retrospective validation study. PLoS Med [Internet]. 2017 Apr 4 [cited 2021 Mar 23];14(4). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5380315/

6 Ethics approval and consent to participate

This project has been approved by the Research Ethics Board of the CHU Sainte-Justine under the reference 2020-2427. In addition, CARTaGENE has obtained ethics approval by the CHU Sainte-Justine under the reference: MP-21-2011-345, 3297. The latest annual ethics renewal was granted on September 13, 2019. This latter approval implies that all participants have given their consent (cartagene.qc.ca/sites/default/files/documents/consent/brochure_en_0505_0.pdf). Written consent was obtained from all the participants.

7 Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

8 Author Contributions

RJ: conceptualization, data curation, formal analysis, investigation, methodology, visualization, writing — original draft, writing — review & editing. NE: writing — review & editing. SCB: writing — review & editing. MCT: writing — review & editing. PB: conceptualization, formal analysis, methodology, project administration, supervision, validation, writing — review & editing.

All authors read and approved the final manuscript.

9 Funding

NE is supported by a clinician scientist award for the Fonds de Recherche du Québec en Santé (FRQS).

10 Acknowledgments

We would like to thank all the CARTaGENE participants for their generous investments in health research. We would also like to thank the RAMQ and the Commission d'accès à l'information (CAI) for their support in obtaining the data.

11 Supplementary Material

Supplementary. Statistical analysis, Table S1 and Table S2

12 Data Availability Statement

The data that support the findings of this study are available from CARTaGENE but restrictions apply to the availability of these data. Data are however available directly from CARTaGENE (http://cartagene.qc.ca; access@cartagene.qc.ca; +1 514-345-2156).

13 Figures

Figure 1. Flow-chart.

Figure 2. Comparison of the screening scenarios. LDCT: low-dose scan. Comparison of the "shift scenario" and the "serial scenario" for one theoretical participant. Percentages are the calculated PLCOm2012 risk. In this example, we considered a threshold of 2% for being screened. For the shift scenario, the PLCOm2012 risk is calculated each year. The participant has a risk higher than 2% during the year 1 and 2, leading to an LDCT each year until the year 7. Its risk remains under 2% until the year 9. The participant had no LDCT during the year 8. The risk was higher than 2% at year 9 and 10, leading to an LDCT per year until the year 15. For the serial scenario, the PLCOm2012 risk score is calculate each 6 years. The participant had a risk score higher than 2% during the year 1. Therefore, he/she had an LDCT per year until year 7. The risk score was calculated at year 8, with a score lower than 2%, leading to the absence of LDCT during the next 6 years. For the Quebec pilot model, these strategies were the same, but no LDCT was made if the participant was outside the age range.

Figure 3. Risk distribution and performance of the PLCOm2012 model (n=11,652). (A) Distribution of the PLCOm2012 model's predictions as a function of cumulative percentage of individuals. (B) Calibration according to the PLCOm2012 model's predictions groups (quartile).

E/O: expected-to-observed cases. (C) Discrimination power of the PLCOm2012 model according to sensitivity and specificity.

14 Tables

Table 1: PLCOm2012 6-year risk prediction accuracy for lung cancer at inclusion (n=11,652)

	PLCOm2012
E/O	0.68 [0.59-0.79]
C-statistic	0.727 [0.679-0.77]
Sensitivity	
Threshold 1.51%	52.3% [44.6-59.8]
Threshold 1.70%	49.4% [41.8-57.1]
Threshold 2.00%	44.9% [37.4-52.6]
Specificity	
Threshold 1.51%	81.6% [80.8-82.3]
Threshold 1.70%	84.3% [83.6-85]
Threshold 2.00%	87.7% [87-88.3]
Positive predictive value	
Threshold 1.51%	4.2% [3.4-5.1]
Threshold 1.70%	4.6% [3.7-5.7]
Threshold 2.00%	5.3% [4.2-6.5]

E/O: expected-to-observed cases

Table 2: Comparison of different computed tomography lung cancer screening inclusion criteria with a shift scenario, between 1998 and 2015 (n=8,938)

	USPSTF	CTFPHC	Quebec Pilot (55-74yrs) + PLCO ≥2%	Quebec Pilot 50-74yrs + PLCO ≥2%	PLCO ≥1.51% (no age criteria)	PLCO ≥1.7% (no age criteria)	PLCO ≥2% (no age criteria)
Total number of participants eligible to be screened*	4445 (49.7%)	2523 (28.2%)	1931 (21.6%)	2045 (22.9%)	2733 (30.6%)	2430 (27.2%)	2064 (23.1%)
Total number of LDCTs	40448	19697	15201	16672	24732	21020	16777
Number of lung cancers detected (n=205)	133 (64.9%)	101 (49.3%)	99 (48.3%)	103 (50.2%)	114 (55.6%)	110 (53.7%)	103 (50.2%)
Number of LDCT for one cancer detected	304.1	195.0	153.5	161.9	216.9	191.1	162.9
Number of participants screened to detect one lung cancer	33.4	25.0	19.5	19.9	24.0	22.1	20.0
Number of LDCT before cancer detection per participant	10.3	8.6	7.7	8.2	9.6	9.1	8.3
Number of LDCT per cancer-free	9.1	7.8	7.9	8.1	9.0	8.6	8.1

	USPSTF	CTFPHC	Quebec Pilot (55-74yrs) + PLCO ≥2%	Quebec Pilot 50-74yrs + PLCO ≥2%	PLCO ≥1.51% (no age criteria)	PLCO ≥1.7% (no age criteria)	PLCO ≥2% (no age criteria)
participants*							
Sensitivity	64.9% [57.9-71.4]	49.3% [42.2-56.3]	48.3% [41.3-55.4]	50.2% [43.2-57.3]	55.6% [48.5-62.5]	53.7% [46.6-60.6]	50.2% [43.2-57.3]
Specificity	50.8% [49.7-51.8]	72.4% [71.4-73.3]	79% [78.2-79.9]	77.8% [76.9-78.6]	70.0% [69.0-71.0]	73.4% [72.5-74.4]	77.5% [76.7-78.4]
Positive predictive value	3.0% [2.5-3.5]	4.2% [3.5-5]	5.1% [4.2-6.2]	5.0% [4.1-6.1]	4.2% [3.5-5.0]	4.5% [3.7-5.4]	5.0% [4.1-6.0]

LDCT: low-dose computed tomography; CTFPHC: Canadian Task Force on Preventive Health Care; PLCOmX%: PLCOm2012 model with X threshold; USPSTF: US Preventive Services Task Force

Shift scenario: we checked eligibility annually. If a participant met the screening criteria, we considered that he/she had an LDCT on the screening date and an LDCT per year during the five next years. If a participant no longer meets the screening criteria, he/she had to complete the remained LDCT.

* only participants with at least one LDCT

Supplementary materials

Statistical analysis

For the PLCOm2012 logistic model, the individual 6-year absolute risk of lung cancer is estimated as:

$$p_i = \frac{exp\left(\alpha + \sum_{j=1}^{n_j} \beta_{ij} X_{ij}\right)}{1 + exp\left(\alpha + \sum_{j=1}^{n_j} \beta_{ij} X_{ij}\right)}$$

where α is a baseline coefficient, β_j is the coefficient associated with the j^{th} risk factor (X_j) and X_{ij} is its value for individual i. Here, n_j is the number of variables included in the PLCOm2012 model. These coefficients and variable can be found in the Tammemägi *et al.* original article [1–3].

1. Tammemägi MC, Katki HA, Hocking WG, Church TR, Caporaso N, Kvale PA, et al. Selection Criteria for Lung-Cancer Screening. N Engl J Med. 2013 Feb 21;368(8):728–36.

2. Lung Cancer Risk Calculators [Internet]. Brock University. [cited 2021 Dec 10]. Available from: https://brocku.ca/lung-cancer-screening-and-risk-prediction/risk-calculators/

3. Tammemägi MC, Church TR, Hocking WG, Silvestri GA, Kvale PA, Riley TL, et al. Evaluation of the Lung Cancer Risks at Which to Screen Ever- and Never-Smokers: Screening Rules Applied to the PLCO and NLST Cohorts. PLOS Medicine. 2014 Dec 2;11(12):e1001764.

Supplementary results

Table S1. Proportion of missing data for the variables included in the PLCOm2012 model

	Number of missing data	
	Six-year risk prediction accuracy for lung cancer at inclusion (n=11,652)	Efficiency of lung cancer screening between 1998 and 2015 (n=8,938)
Age	0	0
Education	26 (0.2%)	15 (0.2%)
Body Mass Index	123 (1.2%)	91 (1.0%)
COPD	74 (0.7%)	799 (8.9%)*
Cancer history	0	0
Family history of lung cancer	328 (3.1%)	258 (2.9%)
Smoking status	0	0
Smoking intensity	1361 (13.0%)	0**
Smoking duration	1456 (13.9%)	0**
Smoking quit time	1066 (10.1%)	0**

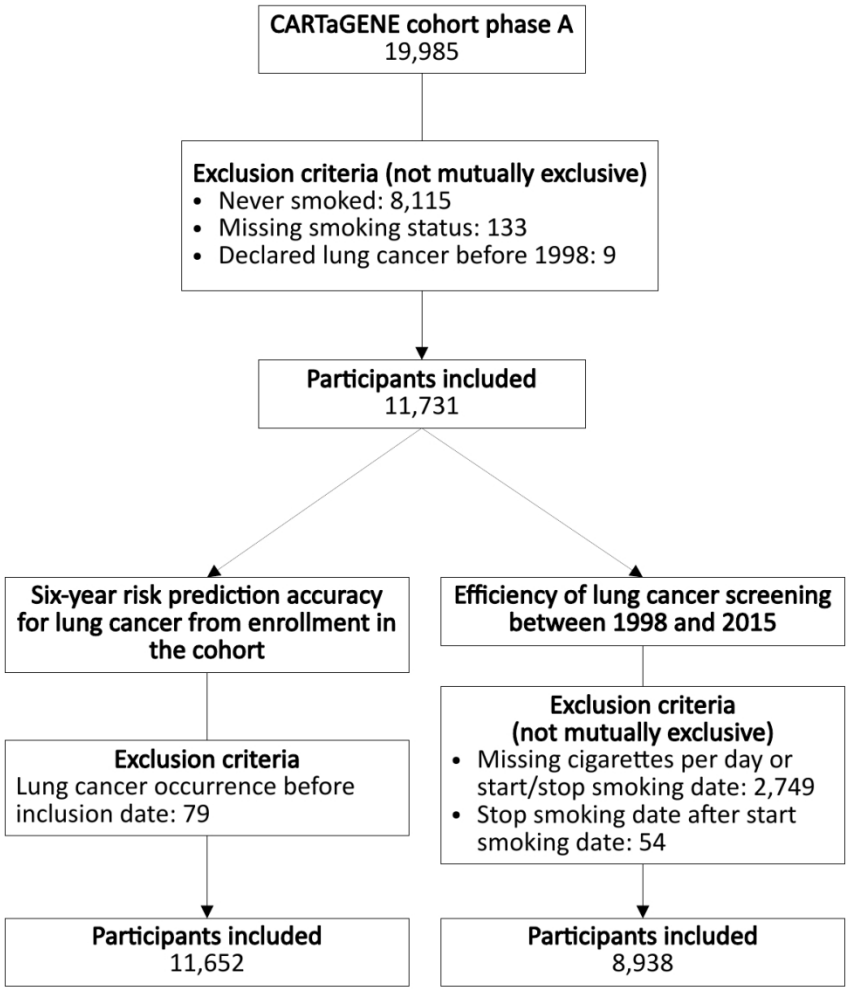
Prediction cohort: 90% of the cohort for validating the models.
*Missing COPD occurrence age
**Participants with missing data were excluded

Table S2. Comparison of different computed tomography lung cancer screening inclusion criteria with a serial scenario, between 1998 and 2015 (n=8938)

	Quebec Pilot (55-74yrs) + PLCO ≥2%	**Quebec Pilot (50-74yrs) + PLCO ≥2%**	**PLCO ≥1.51% (no age criteria)**	**PLCO ≥1.7% (no age criteria)**	**PLCO ≥2% (no age criteria)**
Total number of participants eligible to be screened*	1552 (17.4%)	1554 (17.4%)	2013 (22.5%)	1734 (19.4%)	1408 (15.8%)
Total number of LDCTs	11286	11661	18352	15284	11474
Number of lung cancers detected	78 (38%)	76 (37.1%)	98 (47.8%)	93 (45.4%)	81 (39.5%)
Number of LDCT for one cancer detected	144.7	153.4	187.3	164.3	141.7
Number of participants screened to detect one lung cancer	19.9	20.4	20.5	18.6	16.8
Number of LDCT before cancer detection per participant	7.4	7.7	8.7	8.3	7.0
Number of LDCT per cancer-free participants*	7.3	7.5	9.1	8.8	8.2
Sensitivity	38% [31.4-45.1]	37.1% [30.4-44.1]	47.8% [40.8-54.9]	45.4% [38.4-52.4]	39.5% [32.8-46.6]
Specificity	83.1% [82.3-83.9]	83.1% [82.3-83.9]	78.1% [77.2-78.9]	81.2% [80.4-82]	84.9% [84.1-85.6]

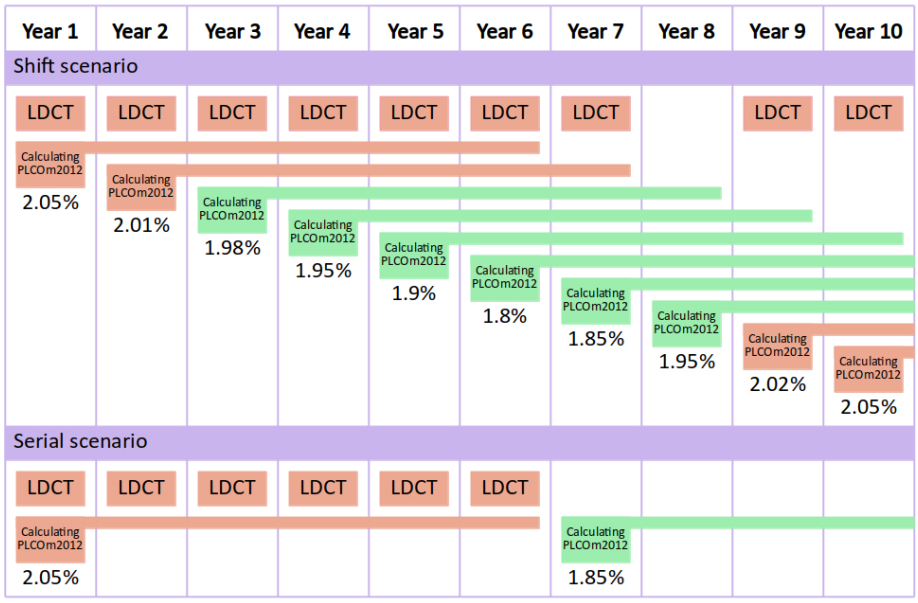
	Quebec Pilot (55-74yrs) + PLCO ≥2%	Quebec Pilot (50-74yrs) + PLCO ≥2%	PLCO ≥1.51% (no age criteria)	PLCO ≥1.7% (no age criteria)	PLCO ≥2% (no age criteria)
Positive predictive value	5.0% [4.0-6.2]	4.9% [3.9-6.1]	4.9% [4-5.9]	5.4% [4.4-6.5]	6.0% [4.8-7.3]

LDCT: low-dose computed tomography; USPSTF: US Preventive Services Task Force
Serial scenario: we checked eligibility each 6 years. If a participant met the inclusion criteria, he/she had a LDCT per year during 6 years. Otherwise, no LDCT was made until the next screening.
* only participants with at least one LDCT

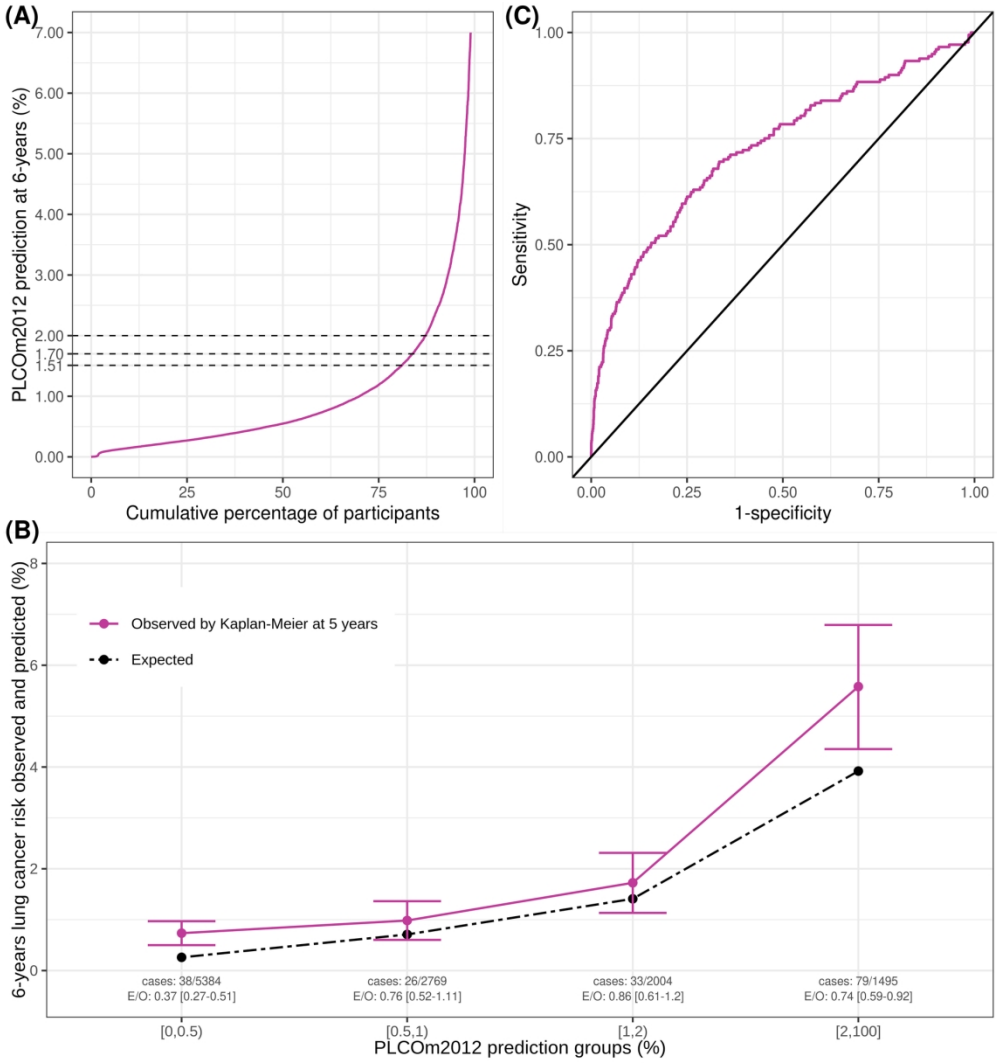


Flow-chart.

215x279mm (150 x 150 DPI)



Comparison of the screening scenarios. LDCT: low-dose scan. Comparison of the "shift scenario" and the "serial scenario" for one theoretical participant. Percentages are the calculated PLCOm2012 risk. In this example, we considered a threshold of 2% for being screened. For the shift scenario, the PLCOm2012 risk is calculated each year. The participant has a risk higher than 2% during the year 1 and 2, leading to an LDCT each year until the year 7. Its risk remains under 2% until the year 9. The participant had no LDCT during the year 8. The risk was higher than 2% at year 9 and 10, leading to an LDCT per year until the year 15. For the serial scenario, the PLCOm2012 risk score is calculate each 6 years. The participant had a risk score higher than 2% during the year 1. Therefore, he/she had an LDCT per year until year 7. The risk score was calculated at year 8, with a score lower than 2%, leading to the absence of LDCT during the next 6 years. For the Quebec pilot model, these strategies were the same, but no LDCT was made if the participant was outside the age range.



Risk distribution and performance of the PLCOm2012 model (n=11,652). (A) Distribution of the PLCOm2012 model's predictions as a function of cumulative percentage of individuals. (B) Calibration according to the PLCOm2012 model's predictions groups (quartile). E/O: expected-to-observed cases. (C) Discrimination power of the PLCOm2012 model according to sensitivity and specificity.

169x179mm (300 x 300 DPI)