

# Evaluation of the accuracy of the PLCO<sub>m2012</sub> 6-year lung cancer risk prediction model among smokers in the CARTaGENE population-based cohort

Rodolphe Jantzen MD, Nicole Ezer MD MPH, Sophie Camilleri-Broët MD PhD, Martin C. Tammemägi PhD, Philippe Broët MD PhD

## Abstract

**Background:** The PLCO<sub>m2012</sub> prediction tool for risk of lung cancer has been proposed for a pilot program for lung cancer screening in Quebec, but has not been validated in this population. We sought to validate PLCO<sub>m2012</sub> in a cohort of Quebec residents, and to determine the hypothetical performance of different screening strategies.

**Methods:** We included smokers without a history of lung cancer from the population-based CARTaGENE cohort. To assess PLCO<sub>m2012</sub> calibration and discrimination, we determined the ratio of expected to observed number of cases, as well as the sensitivity, specificity and positive predictive values of different risk thresholds. To assess the performance of screening strategies if applied between Jan. 1, 1998, and Dec. 31, 2015, we tested different thresholds of the PLCO<sub>m2012</sub> detection of lung cancer over 6 years (1.51%, 1.70% and 2.00%), the criteria of Quebec's pilot program (for people aged 55–74 yr and 50–74 yr) and recommendations from 2021 United States and 2016 Canada guidelines. We assessed shift and serial scenarios of screening, whereby eligibility was assessed annually or every 6 years, respectively.

**Results:** Among 11 652 participants, 176 (1.51%) lung cancers were diagnosed in 6 years. The PLCO<sub>m2012</sub> tool underestimated the number of cases (expected-to-observed ratio 0.68, 95% confidence interval [CI] 0.59–0.79), but the discrimination was good (C-statistic 0.727, 95% CI 0.679–0.770). From a threshold of 1.51% to 2.00%, sensitivities ranged from 52.3% (95% CI 44.6%–59.8%) to 44.9% (95% CI 37.4%–52.6%), specificities ranged from 81.6% (95% CI 80.8%–82.3%) to 87.7% (95% CI 87.0%–88.3%) and positive predictive values ranged from 4.2% (95% CI 3.4%–5.1%) to 5.3% (95% CI 4.2%–6.5%). Overall, 8938 participants had sufficient data to test performance of screening strategies. If eligibility was estimated annually, Quebec pilot criteria would have detected fewer cancers than PLCO<sub>m2012</sub> at a 2.00% threshold (48.3% v. 50.2%) for a similar number of scans per detected cancer. If eligibility was estimated every 6 years, up to 26 fewer lung cancers would have been detected; however, this scenario led to higher positive predictive values (highest for PLCO<sub>m2012</sub> with a 2.00% threshold at 6.0%, 95% CI 4.8%–7.3%).

**Interpretation:** In a cohort of Quebec smokers, the PLCO<sub>m2012</sub> risk prediction tool had good discrimination in detecting lung cancer, but it may be helpful to adjust the intercept to improve calibration. The implementation of risk prediction models in some of the provinces of Canada should be done with caution.

Lung cancer remains the leading cause of cancer-related death in northern America (19.3%) and worldwide (18%).<sup>1</sup> Two large-scaled randomized controlled trials of lung cancer screening, the National Lung Screening Trial and NELSON trial,<sup>2,3</sup> have conclusively shown efficacy (i.e., reduction of lung cancer mortality among high-risk smokers) and cost-effectiveness.<sup>4</sup> As a consequence, screening is now widely supported, but implementation remains limited and varies across countries.<sup>5</sup> Predictive models for lung cancer have been developed with different predicted outcomes (e.g., incidence, death), prediction horizon (e.g., 1 yr, 6 yr) and included risk factors.<sup>5</sup> Among them, the model developed by Tammemägi and colleagues,

**Competing interests:** Nicole Ezer reports funding from the Canadian Institutes of Health Research and Rossy Cancer Network, a speaker fee from GSK, advisory board participation with GSK and receipt of study materials from COVIS Pharma. Martin Tammemägi developed the PLCO<sub>m2012</sub> lung cancer risk prediction models (and related models), used in the current study. To date, he has not received any money for use of the model, nor does he anticipate any payments in the future. No other competing interests were declared.

This article has been peer reviewed.

**Correspondence to:** Rodolphe Jantzen, rodolphe.jantzen@gmail.com; Philippe Broët, philippe.broet@inserm.fr

**CMAJ Open 2023 April 11. DOI:10.9778/cmajo.20210335**

the PLCO<sub>m2012</sub> model,<sup>6</sup> which is used to predict lung cancer at 6 years, showed good discrimination (area under the receiver operating characteristic curve around 0.8). It has been externally validated in different countries<sup>5–8</sup> and, most recently, in the International Lung Screen Trial (Australia; British Columbia, Canada; Hong Kong; the United Kingdom; and Spain) to prospectively identify the best screening strategy between national guidelines and the risk prediction model.<sup>9</sup> Recent findings showed that the PLCO<sub>m2012</sub> was more efficient than the 2013 United States Preventive Services Task Force (USPSTF) criteria for selecting people to enroll into lung cancer screening programs.<sup>10</sup> Moreover, PLCO<sub>m2012</sub> was better than the 2013 USPSTF criteria in terms of sensitivity, deaths averted, screening efficiency and reduction of race and sex disparities.<sup>11–13</sup> In Canada, the Canadian Task Force on Preventive Health Care (CTFPHC)<sup>14</sup> recommends screening for lung cancer using the entry criteria from the National Lung Screening Trial (age 55–74 yr,  $\geq 30$  pack-yr smoking history, smoking quit-time  $< 15$  yr), with low-dose computed tomography scans every year for 3 consecutive screens.

The USPSTF and CTFPHC are both binary criteria, which can lead to the selection of people of too low risk to benefit from screening.<sup>15</sup> In contrast, risk models may be prone to increase the selection of older adults with more comorbidities, which may affect their performance in different jurisdictions. Based on results from cost-effectiveness analyses, Quebec proceeded with using PLCO<sub>m2012</sub> for lung cancer screening, even though this model has not been validated in the Quebec population.<sup>16,17</sup> Therefore, we sought to validate the PLCO<sub>m2012</sub> model among smokers in the CARTaGENE population-based cohort from Quebec to predict the probability of a lung cancer at 6 years. We also sought to compare the efficiency of 7 screening strategies that differed in criteria, frequency of risk score calculation (each year or every 6 yr) and risk score thresholds, if theoretically applied between 1998 and 2015 to our Quebec cohort.

## Methods

### Study population and definition of lung cancer

This study used the CARTaGENE population-based cohort that was recruited in phase A (2009–2010), composed of 19 985 Quebec residents aged 40–69 years.<sup>18</sup> The CARTaGENE cohort consists of adults residing in metropolitan areas, representing 55.7% of the Quebec population (Montreal, Québec, Sherbrooke and Saguenay). Participants were randomly selected to be broadly representative of the population based on provincial health insurance registries (fichier administratif des inscriptions des personnes assurées de la Régie de l'assurance maladie du Québec). Survey design was defined by gender, 2 age groups and forward sortation area (defined by the first 3 digits of postal codes). Participants were excluded if they were not registered in the provincial health insurance registries; if they resided outside the selected regions, in First Nations reserves or in long-term health care facilities; or if they were in prison.

Several strategies were used to obtain adequate response rates and minimize attrition during follow-up phases, including the use of a well-trusted governmental body to contact participants and handle identifying information; the use of systematic methods for contact, scheduling and sending reminders; and financial compensation (\$45). Information packages were first sent by mail, and potential participants were then contacted by telephone to schedule an interview date in one of the clinical assessment sites. Around 35% of the people in the provincial health insurance registries did not have a phone number. Another 13%–15% of the files had incorrect phone numbers. Only files with phone numbers were included in the extraction files as of January 2010 up to October 2010 (the end of recruitment).

Questionnaires at enrolment included data on age, ethnicity, education, body mass index, self-reported history of chronic obstructive pulmonary disease, 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.

We linked participant data with the Quebec administrative health databases from 1998 to 2015 to provide data on cancer diagnoses. We included smokers and people with a history of smoking. We excluded people who had never smoked or had missing smoking data, and those with lung cancer diagnosed before 1998.

As outlined in Tonelli and colleagues,<sup>19</sup> we used administrative data to define incident lung cancer (i.e., people with at least 2 claims in 2 years or 1 hospital admission related to lung cancer; incidence date was the date of first hospital discharge or first claim).

### Study design

Our first objective was to externally validate the PLCO<sub>m2012</sub> model for estimating the 6-year risk of lung cancer from the time of enrolment in the CARTaGENE cohort. The second objective was to determine the hypothetical performance of 7 different screening strategies to detect lung cancers if applied between Jan. 1, 1998, and Dec. 31, 2015. We tested the original PLCO<sub>m2012</sub> model<sup>6</sup> using 3 threshold risks of developing lung cancer over 6 years ( $\geq 1.51\%$ ,  $\geq 1.70\%$  and  $\geq 2.00\%$ ); the 2021 USPSTF criteria (age 50–80 yr, smoker or smoking quit-time  $< 15$  yr,  $\geq 20$  pack-yr smoking history);<sup>20</sup> the 2016 CTFPHC criteria (age 55–74 yr, smoker or smoking quit-time  $< 15$  yr,  $\geq 30$  pack-year smoking history);<sup>14</sup> the Quebec pilot criteria (PLCO<sub>m2012</sub> risk  $\geq 2.00\%$ , age 55–74 yr)<sup>21</sup> and the Quebec pilot criteria with an age range of 50–74 years to test the lower age threshold of the 2021 USPSTF criteria. The risk thresholds for screening selection were based on Pasquini and colleagues;<sup>11</sup> 1.51% has been reported to be a reasonable threshold at which the benefit of mortality reduction benefit for scan over chest radiograph begins, while 1.70% leads to the same number of individuals being selected by the USPSTF criteria; 2.00% was found to be appropriate for use in a pilot study conducted by Ontario Health and Cancer Care Ontario, and is currently used in Ontario for selecting people for scan screening. A summary of the differences

between each strategy can be found in Appendix 1, Supplementary Materials, available at [www.cmajopen.ca/content/11/2/E314/suppl/DC1](http://www.cmajopen.ca/content/11/2/E314/suppl/DC1).

### Statistical analysis

For both objectives, we considered education, family history, smoking status and chronic obstructive pulmonary disease status unchanged after enrolment in the cohort. We replaced missing data for variables in the PLCO<sub>m2012</sub> model by the mean values from Tammemägi and colleagues<sup>6</sup> for continuous variables (age 62 yr, education level 4 [some college education], body mass index 27, duration of smoking 27 yr, smoking quit-time 10 yr) or by the mode, for categorical variables. The proportion of missing data was higher for smoking-related variables such as intensity and duration, but was limited.

For estimating the individual 6-year risk of lung cancer from time of enrolment in the CARTaGENE cohort, we computed the expected-to-observed ratio with 95% confidence intervals (95% CIs), from the sum of the estimated risk (i.e., the number of expected cases) divided by the number of observed cases. We excluded participants with an occurrence of lung cancer before the inclusion date. As some individuals were censored, we obtained the number of observed cases by multiplying the sample size with the Kaplan–Meier estimate of the cumulative lung cancer risk at 6 years. We determined the expected-to-observed ratio in 8 quantiles (octiles) of the risk score. The best calibrated models have an estimate close to 1. We plotted calibration graphs to compare the proportion of observed cases of lung cancer at 6 years in each risk group using a Kaplan–Meier estimator, and the proportion of expected cases (i.e., mean risk). We reported the slope and intercept estimates from logistic regression models (observed outcomes with the logit of the predicted probabilities as the independent variable). We assessed global discrimination by the C-statistic with an inverse probability of censoring weighting estimation of cumulative, time-dependent, receiver operating characteristic curve.<sup>22–24</sup> We calculated sensitivity, specificity and positive predictive value for the 3 threshold risks of developing lung cancer over 6 years ( $\geq 1.51\%$ ,  $\geq 1.70\%$  and  $\geq 2.00\%$ ). We also plotted predictiveness curves (i.e., the risk quantile against the corresponding cumulative proportion of the population with risks below this quantile).

We assessed the hypothetical efficiency of a lung cancer screening strategy as if it has been implemented between Jan. 1, 1998, and Dec. 31, 2015. We excluded participants with missing cigarettes per day, missing start or stop smoking date, or with a stop smoking date before the start smoking date. We considered a lung cancer to be “screen-detected” if the participant was eligible for screening and if a low-dose computed tomography scan would have been theoretically performed in the year before the actual cancer occurrence date. To have at least 1 year post-screening for each participant, the last occurrence of what we considered a screening was in 2014.

We deemed participants as eligible for screening if they 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 PLCO<sub>m2012</sub> model, we determined eligibility based on the shift scenario, whereby we estimated eligibility annually using the PLCO<sub>m2012</sub> thresholds (and added age for the Quebec program), and by the serial scenario, whereby we determined eligibility at 6-year intervals, starting in 1998 for the models using PLCO<sub>m2012</sub> risk criteria and when the participant was aged 50 years or 55 years for the Quebec pilot strategies (Appendix 1, Figure 1).

For each of the 7 screening strategies, we calculated the total number of participants who were theoretically eligible to be screened, the total number of scans that would have been performed, the number of incident lung cancers that would have been detected, the number of scans to be performed to detect 1 lung cancer and the number of participants to be screened to detect 1 lung cancer. We also estimated the number of scans per participant that would have been performed before the detection of the lung cancer and the number of scans per cancer-free participant with at least 1 scan.

We calculated the sensitivity, specificity and positive predictive value. The sensitivity was the probability of being screened in the year before a lung cancer was diagnosed. The specificity was the probability to have no scans per cancer-free year (i.e., the total number of cancer-free years with no scans, divided by the total number of cancer-free years). The positive predictive value was the probability to detect a lung cancer for a participant being screened and having at least 1 scan.

We performed all statistical analyses using R software, version 4.0.<sup>25</sup>

### Ethics approval

This project has been approved by the Research Ethics Board of the CHU Sainte-Justine (no. 2020-2427). In addition, CARTaGENE has obtained ethics approval by the CHU Sainte-Justine (no. MP-21-2011-345, 3297). The latest annual ethics renewal was granted on Sept. 13, 2019. Written consent was obtained from all participants.

### Results

The creation of our study cohort is presented in Figure 1. The cohort characteristics at recruitment can be found in Table 1.

#### Six-year risk prediction accuracy for lung cancer from enrolment in the CARTaGENE cohort

The 11 652 participants included in the cohort used for external validation of the PLCO<sub>m2012</sub> model had a median age of 53.9 (interquartile range [IQR] 48.8–61.0) years at inclusion and a median follow-up time of 5.9 (IQR 5.7–6.0) years. Overall, 176 (1.5%) lung cancers were diagnosed during the 6-year period following enrolment. Using the PLCO<sub>m2012</sub> model, 19.0%, 16.2% and 12.8% of the cohort had a 6-year lung cancer risk that was estimated to be equal or higher than 1.51%, 1.70% and 2.00%, respectively (Figure 2A). The estimated median risk scores for 6-year lung cancer 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.

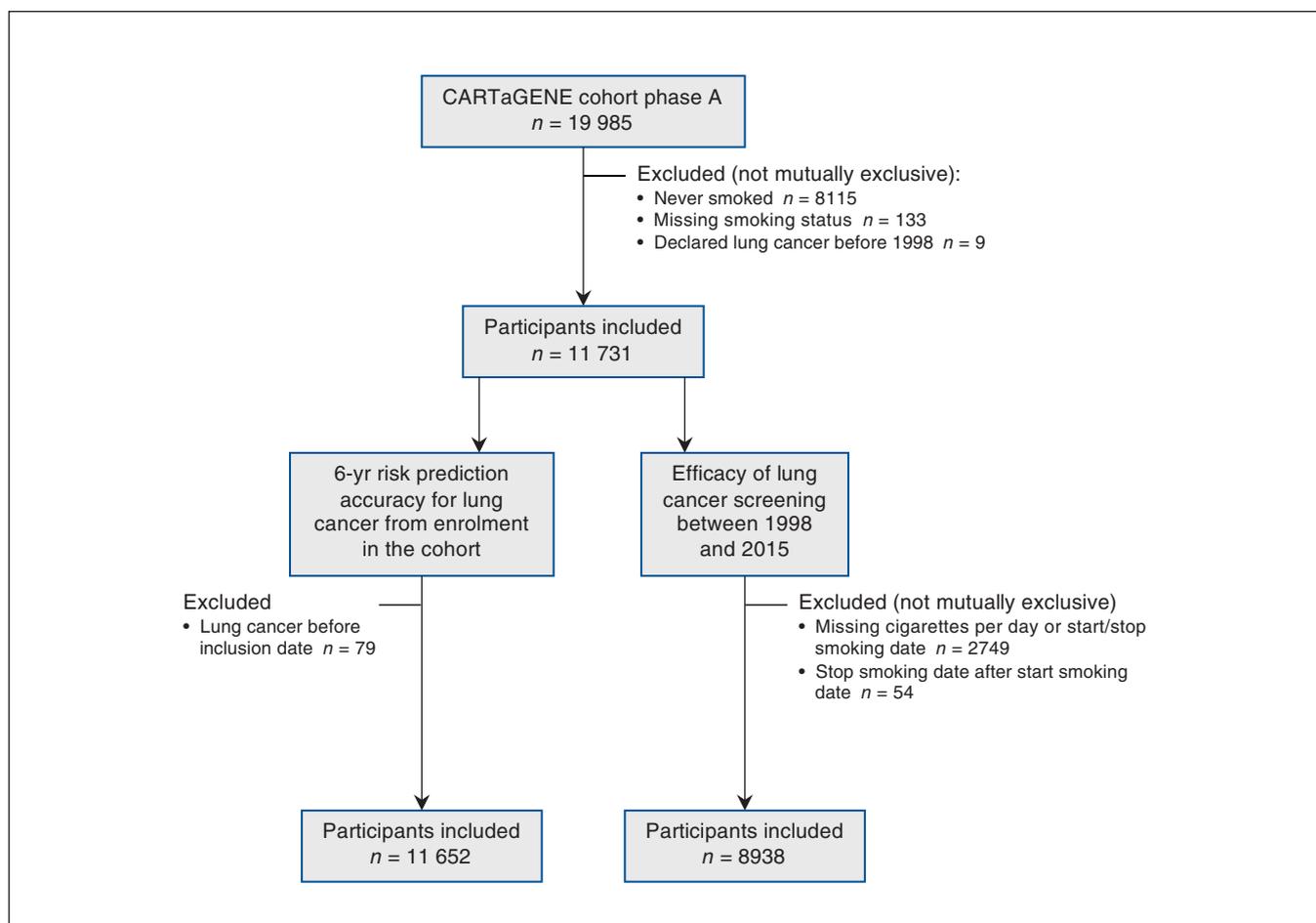


Figure 1: Study flow chart.

The global calibration was 0.68 (95% confidence interval [CI] 0.59–0.79). Expected-to-observed ratios were less than 1 in all risk groups, but this was significant only in the groups with risks of less than 0.27% (0.37, 95% CI 0.27–0.51) and of 2.03% or greater (0.74, 95% CI 0.59–0.92) (Figure 2B). The slope and intercept were of 0.8 (95% CI 0.6 to 0.9) and  $-0.6$  (95% CI  $-1.2$  to 0), respectively. The C-statistic was 0.727 (95% CI 0.679–0.770) (Figure 2C). For the different thresholds, the sensitivity ranged from 44.9% (95% CI 37.4%–52.6%) to 52.3% (95% CI 44.6%–59.8%). The specificity ranged from 81.6% (95% CI 80.8%–82.3%) to 87.7% (95% CI 87.0%–88.3%). The positive predictive value ranged from 4.2% (95% CI 3.4%–5.1%) to 5.3% (95% CI 4.2%–6.5%) (Table 2).

### Hypothetical efficiency of 7 lung cancer screening strategies

Among the 8938 participants included to compare the efficiency of the 7 strategies for lung cancer screening, 205 (2.3%) had a lung cancer between 1998 and 2015.

### Shift scenario

The number of scans that would have been performed ranged from 15 201 (Quebec pilot, age 55–74 yr) to 40 448

(USPSTF), while the number of cancers that would have been detected ranged from 99 (48.3%) (Quebec pilot, age 55–74 yr) to 133 (64.9%) (USPSTF) (Table 3). A 2.00% risk threshold with the  $PLCO_{m2012}$  would have detected more lung cancers than CTFPHC, with fewer scans. The number of lung cancers that would have been detected using the Quebec pilot criteria ( $\geq 2.00\%$  risk and age 55–74 yr) were lower than by using a  $PLCO_{m2012}$  risk threshold of greater than 2.00% alone ( $n = 99, 48.3\%$  v.  $n = 103, 50.2\%$ ) for a similar number of scans performed to detect 1 lung cancer (153.5 v. 162.9). The number of screened participants that would have been needed to detect 1 cancer was the lowest for the Quebec pilot with an age range of 55–74 years (19.5) and highest for the USPSTF (33.4). The USPSTF had the highest sensitivity (64.9%, 95% CI 57.9%–71.4%), and the Quebec pilot had the highest positive predictive value (5.1%, 95% CI 4.2%–6.2%). The results for the Quebec pilot criteria with an age range of 50–74 years were similar to a strategy that used only a 2.00% threshold of  $PLCO_{m2012}$  alone (Table 3).

Using the CTFPHC and USPSTF strategies, 11 and 13 participants would have stopped their screening before the detection of their lung cancer, respectively, as they had stopped smoking for more than 15 years. Their lung

**Table 1: Cohort characteristics at recruitment**

Characteristic	No. (%) of participants*			
	Participants in analysis of 6-yr risk prediction accuracy for lung cancer at inclusion† n = 11 652	Missing	Participants in analysis of efficiency of lung cancer screening between 1998 and 2015 n = 8938	Missing
Age, yr, median (IQR)	53.9 (48.8–61.0)	0	54.1 (49.0–61.0)	0
Gender, woman	5710 (49.0)	0	4260 (47.7)	0
Highest level of education‡		26 (0.2)		15 (0.2)
Some high school	295 (2.4)		235 (2.6)	
High school graduate	3274 (28.2)		2608 (29.2)	
Some college	3846 (33.1)		2980 (33.4)	
College graduate	2073 (17.8)		1534 (17.2)	
Postgraduate or professional degree	2138 (18.4)		1566 (17.6)	
BMI, median (IQR)	26.9 (24.1–30.3)	122 (1.1)	27.1 (24.1–0.5)	91 (1.0)
COPD history	853 (7.4)	74 (0.7)	741 (8.3)	799 (8.9)§
Cancer history	1278 (11.0)	0	1055 (11.8)	0
Family history of lung cancer	1550 (13.7)	321 (2.3)	1251 (14.4)	258 (2.9)
Smoking status		0		0
Daily	2891 (24.8)		2611 (29.2)	
Occasionally	880 (7.6)		480 (5.37)	
Past	7881 (67.6)		5847 (65.4)	
Current no. of cigarettes/d, median (IQR)	18 (13–23)	1352 (11.6)	18 (13–23)	0¶
Smoking duration, yr, median (IQR)	24 (12–34)	1450 (12.4)	25 (14–35)	0¶
Smoking quit time,** yr, median (IQR)	19.9 (10.1–27.9)	1062 (9.1)	19.3 (9.79–27.1)	0¶

Note: BMI = body mass index, COPD = chronic obstructive pulmonary disease, IQR = interquartile range.

\*Unless indicated otherwise.

†Prediction cohort: 90% of the cohort for validating the models.

‡Level 3 (some training after high school) unavailable in our cohort.

§Missing age at COPD occurrence.

¶Participants with missing data were excluded.

\*\*Time in years since a past smoker stopped smoking.

cancers occurred between 3.8 and 14.5 years after their last scans. Among these 13 participants, 4 were detected by the PLCO<sub>m2012</sub> models. None of the participants would have stopped screening before their lung cancer using the other strategies.

### Serial scenario

Among the lung cancers that occurred between 1998 and 2015, the number of cancers that would have been detected using the serial scenario was lower than the shift scenario, ranging from 16 (PLCO<sub>m2012</sub> 1.51%) to 26 (Quebec pilot with ages 50–74 yr) fewer cancers detected. Compared with the shift scenario, the number of screened participants that would have been needed to detect 1 cancer was similar for the Quebec pilot and lower for PLCO<sub>m2012</sub>. The sensitivities were all lower, while the positive predictive values were higher, the highest being the 2.00% threshold of the PLCO<sub>m2012</sub> (6.0%, 95% CI 4.8%–7.3%) (Appendix 1, Table S1 in Supplementary Appendix).

## Interpretation

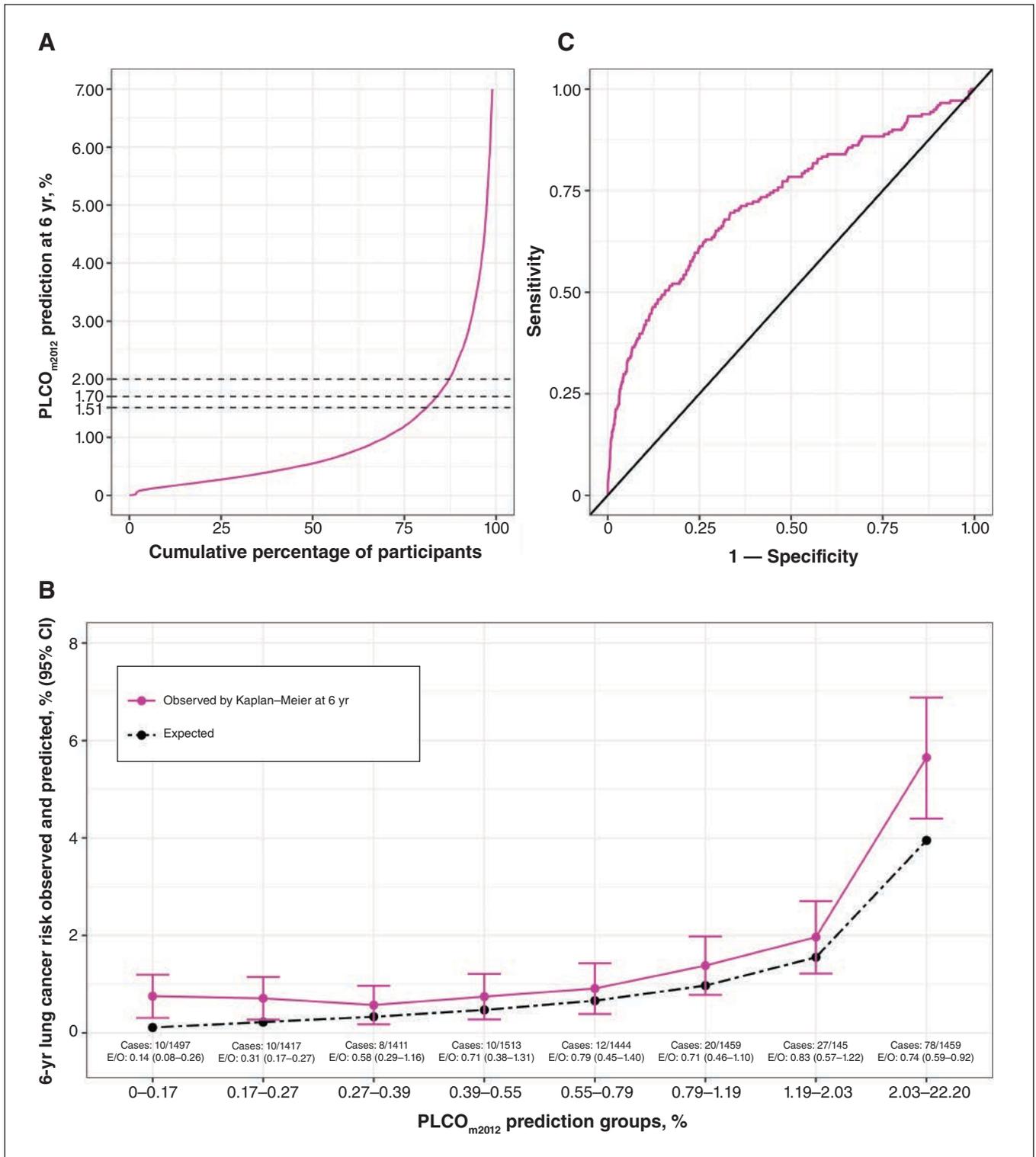
### Six-year risk prediction accuracy for lung cancer from enrolment in the CARTaGENE cohort

We validated the PLCO<sub>m2012</sub> model in the Quebec population; a Quebec pilot study is also prospectively assessing the PLCO<sub>m2012</sub> model for lung cancer screening.<sup>21</sup> In our cohort, for which the rate of new cases of lung cancer was 1.5%, the PLCO<sub>m2012</sub> model underestimated the number of cases. This underestimation can be explained by the higher age-standardized incidence rate of lung cancer in Quebec (106.7, 95% CI 103.3–110.3 cases per 100 000 in 2010) than in the US (88.8, 95% CI 88.3–89.3 cases per 100 000 in 2010), based on information retrieved from national cancer databases.<sup>26–28</sup> This model has 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, given its high incidence of lung cancer, but this should be externally validated.

The risk of lung cancer was overestimated in the UK Biobank, EPIC-UK and Generation Study (incidence < 0.7%, expected-to-observed ratio around 1.3).<sup>8</sup> Compared with the underestimation with a lung cancer incidence of 1.5% in our

cohort, an Australian population-based cohort had an excellent calibration, with an incidence rate of 1.17%.<sup>7</sup>

The discrimination of the PLCO<sub>m2012</sub> was higher in the PLCO cancer screening trial, in UK cohorts and in an



**Figure 2:** Risk distribution and performance of the PLCO<sub>m2012</sub> model ( $n = 11\ 652$ ). Note: CI = confidence interval, E/O = expected-to-observed cases. (A) Distribution of the PLCO<sub>m2012</sub> model's predictions as a function of cumulative percentage of individuals. (B) Calibration according to the PLCO<sub>m2012</sub> model's predictions groups (octile). (C) Discrimination of the PLCO<sub>m2012</sub> model according to sensitivity and specificity.

**Table 2: Six-year risk prediction accuracy for lung cancer at inclusion from the PLCO<sub>m2012</sub> (n = 11 652)**

Variable	Value (95% CI)
Expected-to-observed ratio	0.68 (0.59 to 0.79)
Slope	0.8 (0.6 to 0.9)
Intercept	-0.6 (-1.2 to 0)
C-statistic	0.727 (0.679 to 0.77)
Sensitivity, %	
Threshold 1.51%	52.3 (44.6 to 59.8)
Threshold 1.70%	49.4 (41.8 to 57.1)
Threshold 2.00%	44.9 (37.4 to 52.6)
Specificity, %	
Threshold 1.51%	81.6 (80.8 to 82.3)
Threshold 1.70%	84.3 (83.6 to 85.0)
Threshold 2.00%	87.7 (87.0 to 88.3)
Positive predictive value, %	
Threshold 1.51%	4.2 (3.4 to 5.1)
Threshold 1.70%	4.6 (3.7 to 5.7)
Threshold 2.00%	5.3 (4.2 to 6.5)

Note: CI = confidence interval.

Australian population-based cohort than in our cohort (areas under the receiver operating characteristic curve close to 0.80 v. 0.73 in our cohort).<sup>6-8,29</sup> These differences may be explained by the statistical method used for assessing discrimination (as there were censored data in our study, we used the C-statistic instead of the area under the receiver operating characteristic curve), and how missing data were handled in the UK and Australian cohorts (in these other studies, participants with missing data were excluded or variables with too much missing data were still imputed).

Compared with other studies, we observed lower sensitivity values when using the classical PLCO<sub>m2012</sub> thresholds in the Quebec cohort.<sup>7,11,29</sup> Our positive predictive values were higher,<sup>7</sup> which could be explained by the higher specificities and the higher prevalence of lung cancer in our cohort. This may be owing to smoking exposures in Quebec, which are known to be among the highest in North America.

### Hypothetical efficiency of 7 lung cancer screening strategies

Reassessing eligibility for screening every 6 years instead of annually would lead to far fewer lung cancers being detected and lower sensitivities. However, the higher positive predictive values and the lower cost should be accounted for in public health policies in the absence of a cost-effectiveness

**Table 3: Comparison of different inclusion criteria for lung cancer screening between 1998 and 2015 with a shift scenario (n = 8938)\***

Variable	USPSTF	CTFPHC	Quebec pilot (55-74 yr + PLCO <sub>m2012</sub> ≥ 2%)	Quebec pilot (50-74 yr + PLCO <sub>m2012</sub> ≥ 2%)	PLCO <sub>m2012</sub> ≥ 1.51%	PLCO <sub>m2012</sub> ≥ 1.7%	PLCO <sub>m2012</sub> ≥ 2.0%
Total no. of participants eligible to be screened, n (%) (n = 8938)	4445 (49.7)	2523 (28.2)	1931 (21.6)	2045 (22.9)	2733 (30.6)	2430 (27.2)	2064 (23.1)
Total no. of LDCTs	40 448	19 697	15 201	16 672	24 732	21 020	16 777
No. of lung cancers detected, n (%) (n = 205)	133 (64.9)	101 (49.3)	99 (48.3)	103 (50.2)	114 (55.6)	110 (53.7)	103 (50.2)
No. of LDCTs for 1 cancer detected	304.1	195.0	153.5	161.9	216.9	191.1	162.9
No. of participants screened to detect 1 lung cancer	33.4	25.0	19.5	19.9	24.0	22.1	20.0
No. of LDCTs before cancer detection per participant	10.3	8.6	7.7	8.2	9.6	9.1	8.3
No. of LDCTs per cancer-free participant†	9.1	7.8	7.9	8.1	9.0	8.6	8.1
Sensitivity, % (95% CI)	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, % (95% CI)	73.4% (73.2-73.7)	87.1 (86.9-87.3)	90.0 (89.9-90.2)	89.1 (88.9-89.2)	83.8 (83.6-84.0)	86.2 (86.0-86.4)	89.0 (88.9-89.2)
Positive predictive value, % (95% CI)	3.0% (2.5-3.5)	4.2 (3.5-5.0)	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)

Note: CTFPHC = Canadian Task Force on Preventive Health Care, LDCT = low-dose computed tomography, USPSTF = United States Preventive Services Task Force. \*We checked eligibility annually. If a participant met the screening criteria, we considered that they had an LDCT on the screening date and an LDCT each year during the next 5 years. If a participant no longer met the screening criteria, they had to complete the remained LDCT. †Only participants with at least 1 LDCT.

analysis, as positive predictive values are an important metric for screening policies.

The CTFPHC and 2021 USPSTF criteria seemed less efficient than predictive scores, and in some cases, would have led to screening being stopped before the occurrence of the lung cancer. Therefore, they should not be used in the Quebec population without more precise cost-effectiveness studies.

Decreasing the age limit of the Quebec pilot criteria from 55 to 50 years was equivalent to a 2.00% PLCO<sub>m2012</sub> threshold in terms of the number of cancers that would have been detected, with fewer scans performed and fewer participants screened by the 2.00% threshold of the PLCO<sub>m2012</sub>. However, this last result must be analyzed with participants older than 75 years as the PLCO<sub>m2012</sub> model does not have an age limit and could potentially include this age group. Although the positive predictive value of the 50–74-year age range was slightly lower than that of the 55–74-year age range, decreasing the age threshold would allow the detection of lung cancers among younger people. In the retrospective study by Pasquinelli and colleagues,<sup>11</sup> the 2013 USPSTF had a lower sensitivity than the 2021 USPSTF in our cohort (62.4% v. 64.9%); PLCO<sub>m2012</sub> had better sensitivities than in our cohort (ranging from 60.6% to 70.5%).

More studies are needed to adapt the PLCO<sub>m2012</sub> model to the Quebec population before using it for screening, particularly regarding calibration. Moreover, models such as PLCO<sub>m2012</sub> predict the risk of lung cancer for only a specific period. However, published studies seem to underestimate the frequency with which lung cancer should be screened when using these types of models. Therefore, it is necessary to evaluate different screening scenarios to have a cost-effective screening policy. Finally, a Quebec cohort that includes participants older than 74 years should be used for assessing all of these criteria.

### Limitations

We did not have participants older than 75 years. As the rates of lung cancer are rising among older adults, screening patients older than 75 years may be worthwhile, but it requires further work with cohorts of older adults. 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 a biased estimate. However, since the population is coming from metropolitan areas with a quite homogeneous public system of health care delivery, this should not represent a major issue. Some self-reported variables were only available before the inclusion date and, therefore, the incidence during follow-up could not be retrieved (e.g., chronic obstructive pulmonary disease). However, the time horizon of 6 years considered for this study means that this problem should potentially occur in a very low proportion. We had imputed missing data in the PLCO<sub>m2012</sub> model, particularly regarding smoking-related variables, but the proportion was lower than in other large cohort studies.<sup>7,8</sup> The CARTaGENE cohort is broadly representative of the Quebec population, but the participants were

generally more educated, with a slight over-representation of people of racial and ethnic minority groups, and not all the regions of Quebec were included. However, since the objective was to evaluate screening strategies for lung cancer, the CARTaGENE cohort is nevertheless well suited since it gathers a target population that would be easily accessible for lung cancer screening (i.e., metropolitan centres and educated people). It is worth noting that not accounting for competing events such as smoking-related death may be a limitation of our work when computing the number of events in our study. However, the study's time interval was short and the number of deaths was limited. Finally, we considered that lung cancers were detected if a scan was made 1 year before the cancer's incidence, which was arbitrary but similar across the investigated scenarios.

### Conclusion

In a cohort of Quebec smokers, the PLCO<sub>m2012</sub> risk prediction tool had good discrimination in detecting lung cancer, but it may be helpful to adjust the intercept to improve calibration. In addition, our findings support that the estimation of lung cancer risk and screening eligibility should be done every 6 years with a 2.00% threshold of the PLCO<sub>m2012</sub>. Lowering the onset age of screening to 50 years from 55 years may be considered but would require further cost-effectiveness analyses. Finally, the CTFPHC criteria seemed less efficient than predictive scores; therefore, our results indicate that these criteria should not be used in Quebec.

### References

1. Lung [fact sheet]. Lyon (France): International Agency for Research on Cancer, World Health Organization; 2020. Available: <https://gco.iarc.fr/today/data/factsheets/cancers/15-Lung-fact-sheet.pdf> (accessed 2021 Apr. 22).
2. National Lung Screening Trial Research Team; 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.
3. Zhao YR, Xie X, de Koning HJ, et al. NELSON lung cancer screening study. *Cancer Imaging* 2011;11:S79-84.
4. Goffin JR, Flanagan WM, Miller AB, et al. Cost-effectiveness of Lung Cancer Screening in Canada. *JAMA Oncol* 2015;1:807-13.
5. Sands J, Tammemägi MC, Couraud S, et al. Lung screening benefits and challenges: a review of the data and outline for implementation. *J Thorac Oncol* 2021;16:37-53.
6. Tammemägi MC, Katki HA, Hocking WG, et al. Selection criteria for lung-cancer screening. *N Engl J Med* 2013;368:728-36.
7. Weber M, Yap S, Goldsbury D, et al. Identifying high risk individuals for targeted lung cancer screening: independent validation of the PLCO<sub>m2012</sub> risk prediction tool. *Int J Cancer* 2017;141:242-53.
8. 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.
9. Lim KP, Marshall H, Tammemägi M, et al.; ILST (International Lung Screening Trial) Investigator Consortium. Protocol and rationale for the International Lung Screening Trial. *Ann Am Thorac Soc* 2020;17:503-12.
10. Tammemägi MC, Ruparel M, Tremblay A, et al. USPSTF2013 versus PLCO<sub>m2012</sub> lung cancer screening eligibility criteria (International Lung Screening Trial): interim analysis of a prospective cohort study. *Lancet Oncol* 2022;23:138-48.
11. Pasquinelli MM, Tammemägi MC, Kovitz KL, et al. Risk prediction model versus United States Preventive Services Task Force lung cancer screening eligibility criteria: reducing race disparities. *J Thorac Oncol* 2020;15:1738-47.
12. Ten Haaf K, Bastani M, Cao P, et al. A comparative modeling analysis of risk-based lung cancer screening strategies. *J Natl Cancer Inst* 2020;112:466-79.
13. Tammemägi MC, Church TR, Hocking WG, 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 Med* 2014;11:e1001764.

14. Canadian Task Force on Preventive Health Care. Recommendations on screening for lung cancer. *CMAJ* 2016;188:425-32.
15. Tammemägi MC. Selecting lung cancer screenees using risk prediction models: where do we go from here. *Transl Lung Cancer Res* 2018;7:243-53. Available: <https://tcr.amegroupp.com/article/view/21997> (accessed 2021 Apr. 29).
16. Darling GE, Tammemägi MC, Schmidt H, et al. Organized lung cancer screening pilot: informing a province-wide program in Ontario, Canada. *Ann Thorac Surg* 2021;111:1805-11.
17. Tammemägi MC, Darling GE, Schmidt H, et al. Selection of individuals for lung cancer screening based on risk prediction model performance and economic factors: the Ontario experience. *Lung Cancer* 2021;156:31-40.
18. Awadalla P, Boileau C, Payette Y, et al.; CARTaGENE Project. Cohort profile of the CARTaGENE study: Quebec's population-based biobank for public health and personalized genomics. *Int J Epidemiol* 2013;42:1285-99.
19. Tonelli M, Wiebe N, Fortin M, et al.; Alberta Kidney Disease Network. Methods for identifying 30 chronic conditions: application to administrative data. *BMC Med Inform Decis Mak* 2015;15:31.
20. US Preventive Services Task Force; Krist AH, Davidson KW, Mangione CM, et al. Screening for lung cancer: US Preventive Services Task Force recommendation statement. *JAMA* 2021;325:962-70.
21. Lung cancer screening in Canada: environmental scan (2019–2020). Toronto: Canadian Partnership Against Cancer; updated 2021 Jan. 13:1-41.
22. Uno H, Cai T, Tian L, et al. Evaluating prediction rules for *t*-year survivors with censored regression models. *JASA* 2007;102:527–37.
23. Blanche P, Latouche A, Viallon V. Time-dependent AUC with right-censored data: a survey study. *ArXiv* 2012 Oct 25. Available: <http://arxiv.org/abs/1210.6805> (accessed 2017 Sept. 26).
24. Blanche P, Dartigues JF, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Stat Med* 2013;32:5381–97.
25. R Core Team: A language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing. Available: <https://www.R-project.org/> (accessed 2022 Oct. 25).
26. Table 13-10-0111-01: Number and rates of new cases of primary cancer, by cancer type, age group and sex. Ottawa: Statistics Canada. Available: <https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1310011101> (accessed 2019 Nov. 10).
27. Table 17-10-0005-01: Population estimates on July 1st, by age and sex. Ottawa: Statistics Canada; Available: <https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1710000501> (accessed 2019 Nov. 10).
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; 2021. Available: <https://www.cdc.gov/cancer/dataviz> (accessed 2021 Sept. 16).
29. Ten Haaf K, Jeon J, Tammemägi MC, et al. Risk prediction models for selection of lung cancer screening candidates: a retrospective validation study. *PLoS Med* 2017;14:e1002277.

**Affiliations:** CARTaGENE (Jantzen, Broët), Research Centre, CHU Sainte-Justine; Université de Montréal (Jantzen, Broët); Département de médecine sociale et préventive (Broët), École de santé publique de l'Université de Montréal, Université de Montréal; Departments of Medicine (Ezer) and of Pathology (Camilleri-Broët), McGill University, Montréal, Que.; Prevention and Cancer Control (Tammemägi), Ontario Health (Cancer Care Ontario), Toronto, Ont.; Department of Health Sciences (Tammemägi), Brock University, St. Catharines, Ont.; Department of Public Health (Broët), Faculty of Medicine; Centre de recherche en épidémiologie et santé des populations and INSERM (Broët), Université Paris-Saclay; Assistance Publique-Hôpitaux de Paris (Broët), Hôpitaux Universitaires Paris-Sud, Hôpital Paul Brousse, Villejuif, France

**Contributors:** Rodolphe Jantzen, Philippe Broët and Martin Tammemägi contributed to the conception and design of the work. Rodolphe Jantzen and Philippe Broët contributed to data acquisition and analysis. Rodolphe Jantzen, Philippe Broët, Martin Tammemägi, Nicole Ezer and Sophie Camilleri-Broët contributed to data interpretation and writing. All of the authors revised it critically for important intellectual content, gave final approval of the version to be published and agreed to be accountable for all aspects of the work.

**Funding:** Nicole Ezer is supported by a clinician–scientist award for the Fonds de recherche du Québec – Santé.

**Content licence:** This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY-NC-ND 4.0) licence, which permits use, distribution and reproduction in any medium, provided that the original publication is properly cited, the use is noncommercial (i.e., research or educational use), and no modifications or adaptations are made. See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>

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

**Acknowledgements:** The authors would like to thank all of the CARTaGENE participants for their generous investments in health research. They would also like to thank the Régie de l'assurance maladie du Québec (RAMQ) and the Commission d'accès à l'information (CAI) for their support in obtaining the data.

**Supplemental information:** For reviewer comments and the original submission of this manuscript, please see [www.cmajopen.ca/content/11/2/E314/suppl/DC1](http://www.cmajopen.ca/content/11/2/E314/suppl/DC1).