

# Predictive Accuracy of the PanCan Lung Cancer Risk Prediction Model -External Validation based on CT from the Danish Lung Cancer Screening Trial

Mathilde M. Winkler Wille · Sarah J. van Riel ·  
Zaigham Saghir · Asger Dirksen · Jesper Holst Pedersen ·  
Colin Jacobs · Laura Hohwü Thomsen · Ernst Th. Scholten ·  
Lene T. Skovgaard · Bram van Ginneken

Received: 18 November 2014 / Revised: 18 February 2015 / Accepted: 23 February 2015 / Published online: 13 March 2015  
© European Society of Radiology 2015

## Abstract

**Objectives** Lung cancer risk models should be externally validated to test generalizability and clinical usefulness. The Danish Lung Cancer Screening Trial (DLCST) is a population-based prospective cohort study, used to assess the discriminative performances of the PanCan models.

**Methods** From the DLCST database, 1,152 nodules from 718 participants were included. Parsimonious and full PanCan risk prediction models were applied to DLCST data, and also coefficients of the model were recalculated using DLCST data. Receiver operating characteristics (ROC) curves and area under the curve (AUC) were used to evaluate risk discrimination.

**Results** AUCs of 0.826–0.870 were found for DLCST data based on PanCan risk prediction models. In the DLCST, age and family history were significant predictors ( $p=0.001$  and  $p=0.013$ ). Female sex was not confirmed to be associated with higher risk of lung cancer; in fact opposing effects of sex were observed in the two cohorts. Thus, female sex appeared to lower the risk ( $p=0.047$  and  $p=0.040$ ) in the DLCST.

**Conclusions** High risk discrimination was validated in the DLCST cohort, mainly determined by nodule size. Age and family history of lung cancer were significant predictors and could be included in the parsimonious model. Sex appears to be a less useful predictor.

## Key points

- High accuracy in logistic modelling for lung cancer risk stratification of nodules.
- Lung cancer risk prediction is primarily based on size of pulmonary nodules.
- Nodule spiculation, age and family history of lung cancer are significant predictors.
- Sex does not appear to be a useful risk predictor.

M. M. Winkler Wille (✉) · A. Dirksen  
Department of Respiratory Medicine, Gentofte Hospital,  
Kildegårdsvej 28, Opg.1D, st.th., DK-2900 Hellerup, Denmark  
e-mail: mathilde.winkler@gmail.com

S. J. van Riel · C. Jacobs · E. T. Scholten · B. van Ginneken  
Department of Radiology and Nuclear Medicine,  
Radboud University Medical Center, Nijmegen, The Netherlands

Z. Saghir  
Department of Respiratory Medicine, Herlev Hospital,  
Herlev, Denmark

J. H. Pedersen  
Department of Thoracic Surgery, Rigshospitalet,  
Copenhagen University Hospital, København Ø, Denmark

L. H. Thomsen  
Department of Respiratory Medicine, Hvidovre Hospital,  
Hvidovre, Denmark

L. T. Skovgaard  
Department of Biostatistics, University of Copenhagen,  
København Ø, Denmark

**Keywords** Lung cancer screening · Diagnostic imaging ·  
Computed tomography · Risk · Solitary pulmonary nodules

## Abbreviations

AUC	Area under the curve
BCCA	British Columbia Cancer Agency
CT	Computed tomography
DLCST	Danish Lung Cancer Screening Trial
NLST	National Lung Screening Trial
OR	Odds ratio

PanCan	Pan-Canadian Early Detection of Lung Cancer Study
ROC	Receiver operating characteristics
SD	Standard deviation

## Introduction

Lung cancer is the primary cause of cancer-related death in the world [1]; thus, lung cancer screening trials have been performed widely in an attempt to intervene in early stage disease—hopefully resulting in reduced mortality. The NLST (National Lung Screening Trial) showed a 20 % reduction in lung cancer mortality as well as 6.7 % decrease in all-cause mortality with low-dose computed tomography (CT) screening versus conventional chest radiography [2]. However, the benefit-to-harm ratio is still not fully investigated and false positive screens remain a worrying issue [3]. Targeting lung cancer screening and maximising screening benefits are the focus of future research. In this respect, selecting high-risk individuals and identifying high-risk pulmonary nodules are crucial elements of any lung cancer screening program. Various risk prediction models using patient characteristics as well as clinical risk factors have been suggested, some focusing on pre-screen or early-screen risk for selection of high-risk individuals (where no nodules have been found yet) [4–8], others focus on risk of malignancy of pulmonary nodules detected in screening [9–12].

In order for a model to prove clinically useful, it must prove not only to perform well in the original cohort, from which it was developed, but also on external data sets, thus documenting generalizability. Therefore, validation of suggested risk prediction models is essential.

The recently proposed risk prediction models exploring the probability of cancer in pulmonary nodules detected on first screening CT suggested by McWilliams et al. [12] were performed on the cohort originating from the Pan-Canadian Early Detection of Lung Cancer Study (PanCan). For validation, data from participants from chemoprevention trials from the British Columbia Cancer Agency (BCCA) were used. Parsimonious and full logistic multivariable regression models were performed. The parsimonious model included variables with *p* less than 0.05: sex, nodule size and nodule location. In the full model, variables with *p* values less than 0.25 and greater than 0.05 were added: age, family history of lung cancer, visually assessed emphysema, nodule type (nonsolid or with ground glass opacity, part-solid or solid) and nodule count per scan. Both the parsimonious and the full model were performed with and without the inclusion of the variable spiculation, because this variable was not available from the BCCA data set. In the study, receiver operating characteristics

(ROC) curves with areas under the curve (AUC) of at least 0.90 were achieved.

The Danish Lung Cancer Screening Trial (DLCST) is a population-based prospective randomised controlled trial with extended data on patient characteristics and both nodule characteristics and visual evaluations of emphysema [13]. In this study, we assess the discriminative performance of the PanCan model using the DLCST cohort.

## Method

### Study population

We used data from the DLCST, in which participants between the ages of 50 and 70 years with a smoking history of at least 20 pack-years, and a lung function of at least 30 % of the predicted, were randomised to either five annual low-dose CT examinations or annual control visits without imaging. Participants were either continuous smokers or ex-smokers who had quit after the age of 50 years and no more than 10 years prior to entering the study. A detailed description of the study has been previously published [14].

The DLCST was approved by the Ethics Committee of Copenhagen County and fully funded by the Danish Ministry of Interior and Health. Approval of data management in the trial was obtained from the Danish Data Protection Agency. The trial is registered in the ClinicalTrials.gov Protocol Registration System (identification no. NCT00496977). All participants provided written informed consent.

### Imaging

The screening group was examined on CT annually over a period of 5 years, using a multi-slice CT system (16-row Philips Mx 8000, Philips Medical Systems). Examinations were performed supine after full inspiration using a low dose technique (120 kV and 40 mAs) with the following specifications: section collimation 16×0.75 mm, pitch 1.5 and rotation time 0.5 s. Participants were instructed to first hyperventilate three times and then inhale maximally and hold their breath during imaging. Images were reconstructed with two kernels: thick (3 mm) and thin (1 mm) slice thicknesses using soft and hard algorithms (kernel C and D), respectively. Nodule analyses and visual assessments were performed on thin slices.

### Nodule data

In this study, participants randomised to the screening arm of DLCST with at least one nodule without benign calcification pattern were included. Benign calcification pattern was defined as central, laminated, popcorn or diffuse calcification.

Two experienced chest radiologists (KB and HH) were responsible for the initial assessments of the images during the course of the screening, in which nodule size was manually measured and given as an average of the two observations. A nodule diameter of 3 mm was considered the lower limit of a positive finding in the initial evaluation in DLCST. One experienced chest radiologist (ES), blinded to diagnoses of lung cancer, recorded spiculation and patterns of benign or potentially malignant calcification, and categorised the nodules according to type: perifissural, solid, part-solid or non-solid (pure ground glass). All nodules found throughout the study period were included with the image on which they were first seen. Perifissural nodules were not included in the prediction analysis using PanCan model coefficients, as no coefficient was stated for this nodule subtype owing to the very low risk of malignancy. In this study, nodule count during the whole period of observation was used as opposed to nodule count per scan used in the PanCan study; the DLCST data registration only allowed for this analysis. One nodule which was classified as a benign calcification turned out to be malignant. This nodule was removed from the analysis as a result of the benign classification by the blinded radiologist (ES).

#### Data on emphysema

Baseline study images were evaluated for emphysema by one observer (MW, radiology resident with 3 years of chest radiology experience) blinded to study identifiers as well as presence or absence of lung cancer in the individual participant. Thus, images were selected so no lung cancer findings were revealed, and in case a cancer was present in the baseline scan, the second round scan, in which the cancer had been removed, was chosen instead. Consequently, no lung cancers were shown to the observer, who recorded emphysema as present or not present. If scans without nodules were unavailable, this participant was not included in the emphysema assessment. A detailed description of the method for assessment of emphysema has been previously published [13].

#### Statistical analysis

Basic comparative statistics were performed with the use of Student's *t* test for continuous data and Fisher's exact test for categorical data.

To estimate the lung-cancer-risk predictive ability of the PanCan models, risks were calculated with coefficients and intercepts in accordance with the PanCan risk models, using both the parsimonious and full models with spiculation (1b and 2b). For nodule size, transformations were performed as a result of non-linearity in the relationship between size and risk of lung cancer [12]. The ability of the PanCan risk prediction models to separate persons in the DLCST cohort who developed lung cancer from those who did not was assessed by

measuring discriminative accuracy with the use of ROC and AUC. Secondly, multivariable logistic regression analyses were performed using DLCST data and covariates from the PanCan parsimonious and full models, thereby estimating new regression coefficients for comparison. Because some participants had more than one nodule, the variances of effect estimates were adjusted for correlated responses (clustering of data within persons, Huber–White method). ROC and AUC were used to evaluate risk prediction. Difference between AUCs was tested by use of paired bootstrapping based on 1,000 bootstrapped samples. Statistical programming and figures were performed using R, version 3.1.1, packages rms, pROC and ROCR.

## Results

In DLCST, 823 persons were diagnosed with 1,385 nodules of which 233 nodules were classified as benign calcifications and excluded, leaving a total of 718 persons and 1,152 nodules to be included in the analyses.

Table 1 shows nodule characteristics by lung cancer status, in PanCan, BCCA and DLCST, respectively. The smallest nodule diameter in DLCST was 3 mm; hence, nodules are generally larger than in the PanCan and BCCA cohorts, in which nodules with a diameter of only 1 mm were included. Mean diameters differ markedly from median diameters, and the distributions of nodule sizes are skewed. Therefore, mean nodule sizes and standard deviations are of limited interpretive value.

Malignancy was significantly related to nodule type (Table 1); thus, perifissural nodules had a very low risk (OR 0.1,  $p=0.017$ ), and part-solid nodules were associated with increased risk (OR 2.9,  $p=0.009$ ).

Included participants had a mean age of 59.0 years (SD 4.9); participants with cancer were slightly older than participants without cancer (mean age 61.6 (SD 4.9) versus 58.8 (SD 4.8) years,  $p<0.001$ ).

There were 338 women (47.1 %) and 380 (52.9 %) men included in the study. Of these, 66 were diagnosed with lung cancer, six participants had more than one malignant nodule. Lung cancer tended to be less frequent in women compared to men (OR 0.71,  $p=0.205$ ), and emphysema was a co-finding of 36.9 % of benign nodules and of 38.7 % of malignant nodules (OR 1.1,  $p=0.788$ ). Thirteen participants (four with lung cancer and nine without) with nodules included in this study lack emphysema assessment and were thus excluded from the full model logistic regression.

Tables 2 and 3 show results from the parsimonious and full models, respectively, comparing PanCan and DLCST results. In PanCan, female sex implied a higher risk (parsimonious model: OR 1.91 (CI 1.19–3.07),  $p=0.008$ ; full model: OR 1.82 (CI 1.12–2.97),  $p=0.02$ ), whereas in DLCST, female

**Table 1** Nodule characteristics in the three different cohorts

Variable	PanCan development cohort <sup>b</sup>			BCCA validation cohort <sup>b</sup>			DLCST validation cohort			OR	p value
	Benign nodule N=6,906	Lung cancer N=102	Total N=7,008	Benign nodule N=4,979	Lung cancer N=42	Total N=5,021	Benign nodule N=1,086	Lung cancer N=66	Total N=1,152		
<b>Nodule size (mm)</b>											
Mean <sup>a</sup>	4.1±3.1	15.7–12.2	4.3±3.7	3.6±2.2	13.9±7.4	3.7±2.5	6.9±5.0	14.4±12.2	7.3±5.9		<0.001 <sup>c</sup>
Median	3.4	13	3.5	3	12.8	3	6.0	11.0	6.0		
Range	1–70	2–86	1–86	1–45	3–45	1–45	3–90	4–93	3–93		
Interquartile range	2.7–5.0	8–19	2.8–5.0	2.0–4.5	9–18	2.0–4.8	4.0–8.0	8.0–17	4.0–8.0		
<b>Nodule type</b>											
Perifissural	70	0	6,989	501	0	5,021	113	1	1,152		0.004 <sup>d</sup>
Non-solid (ground glass)	1,084	21	1,105	461	6	467	116	8	114		0.1
Part-solid	283	20	303	35	10	45	56	9	124		1.2
Solid	5,450	61	5,511	3,982	26	4,008	801	48	849		2.9
<b>Nodule location</b>											
Left lower	1,116	12	6,964	852	11	863	221	12	233		0.030 <sup>e</sup>
Left upper	1,581	31	1,612	1,277	8	1,285	178	15	193		
Right lower	1,249	21	1,270	829	1	830	262	10	272		
Right middle	685	2	687	503	1	504	99	4	103		
Right upper	2,232	35	2,267	1,518	21	1,539	326	25	351		
<b>Nodule count</b>											
Mean	6.2±4.0	4.8±3.5	6.2±4.0			10.3±9.8	2.5±1.7	2.1±1.7	2.4±1.7		0.115 <sup>c</sup>
Median	5	4	5			7	2.0	1.0	2.0		
Range	1–31	1–19	1–31			1–60	1–10	1–8	1–10		
Interquartile range	3–9	2–6	3–9			4–13	1.0–3.0	1.0–3.0	1.0–3.0		
<b>Spiculation</b>											
No	6,739	72	6,811				1,058	50	1,108		<0.001 <sup>d</sup>
Yes	167	30	197				28	16	44		12.0

<sup>a</sup> Mean±standard deviation (SD)<sup>b</sup> McWilliams et al. (NEJM 369:910–919, 2013)<sup>c</sup> Student's *t* test with unequal variance<sup>d</sup> Fisher's exact test<sup>e</sup> Regarding nodule location, Fisher's exact test on upper versus middle/lower lobes: OR 1.8, *p*=0.030

**Table 2** PanCan parsimonious prediction model 1b for the probability of lung cancer

Predictor variables	PanCan parsimonious model 1b					
	PanCan data			DLCST data		
	Odds ratio (95 % CI)	<i>P</i> value	Coefficient	Odds ratio (95 % CI)	<i>P</i> value	Coefficient
Sex, female vs. male	1.91 (1.19–3.07)	0.008	0.6467	0.55 (0.31–0.96)	0.047	−0.6013
Nodule size <sup>a</sup>		<0.001	−5.5537		<0.001	−4.1209
Nodule location, upper vs. middle/lower lobe	1.82 (1.12–2.98)	0.02	0.6009	2.24 (1.27–3.96)	0.007	0.8081
Spiculation, yes vs. no	2.54 (1.45–4.43)	0.001	0.9309	3.77 (1.72–8.30)	0.002	1.3267
Model constant			−6.6144			−4.9133

<sup>a</sup> 4 mm subtracted from actual nodule size. This is done in accordance with PanCan original models to centre on PanCan mean nodule size

sex tended to lower the risk (parsimonious model: OR 0.55 (CI 0.31–0.96), *p*=0.047; full model: OR 0.49 (CI 0.26–0.91), *p*=0.040). Furthermore, in DLCST age (OR 1.10 (CI 1.03–1.17), *p*=0.001) and family history (OR 2.61 (CI 1.37–4.98), *p*=0.013) were significant predictor variables in the full model (Table 3). Spiculation is a major predictor of malignancy in DLCST (parsimonious model: OR 3.77 (CI 1.72–8.30), *p*=0.002; full model: OR 3.40 (CI 1.36–8.46), *p*=0.013); and nodule size is by far the most important determinant in both the parsimonious and full models (regression coefficients −4.1909 and −3.8075, respectively, *p*<0.001). ROC curve analysis using exclusively nodule size as the determinant variable, with DLCST coefficients, resulted in an AUC of 0.829.

For the parsimonious model with DLCST data (Fig. 1), AUCs were 0.826 and 0.853 using PanCan

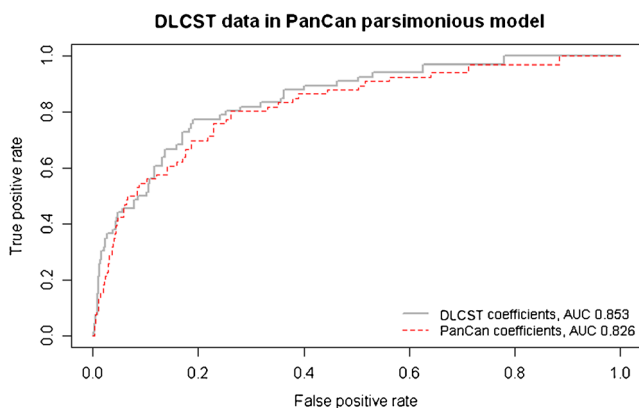
and DLCST coefficients, respectively; and for the full model (Figs. 2 and 3), AUCs were 0.834 and 0.870 using PanCan and DLCST coefficients, respectively. Using DLCST coefficients, AUC of the full model was not significantly larger than AUC of the parsimonious model (AUC 0.870 vs. 0.853, *p*=0.064). AUC of the full model is significantly larger than AUC if nodule size is the only determinant (AUC 0.870 vs. 0.829, *p*=0.015), but AUC of the parsimonious model was not significantly larger than AUC if nodule size is the only determinant (AUC 0.853 vs. 0.829, *p*=0.065).

The non-linear transformation of nodule size used in the PanCan study ( $1/\sqrt{\text{nodule diameter}}$ ) resulted in the same AUCs as did use of the transformation  $1/\log(\text{nodule diameter})$ .

**Table 3** PanCan full prediction model 2b for the probability of lung cancer

Predictor variables	PanCan full model 2b					
	PanCan data			DLCST data		
	Odds ratio (95 % CI)	<i>P</i> value	Coefficient	Odds ratio (95 % CI)	<i>P</i> value	Coefficient
Age, per year <sup>a</sup>	1.03 (0.99–1.07)	0.16	0.0287	1.10 (1.03–1.17)	0.001	0.0949
Sex, female vs. male	1.82 (1.12–2.97)	0.02	0.6011	0.49 (0.26–0.91)	0.040	−0.7134
Family history of lung cancer, yes vs. no	1.34 (0.83–2.17)	0.23	0.2961	2.61 (1.37–4.98)	0.013	0.9582
Emphysema, yes vs. no	1.34 (0.78–2.33)	0.29	0.2953	0.74 (0.40–1.36)	0.382	−0.3024
Nodule size <sup>a</sup>		<0.001	−5.3854		<0.001	−3.8075
Nodule type						
Non-solid (ground glass)	0.88 (0.48–1.62)	0.68	−0.1276	0.95 (0.39–2.28)	0.910	−0.0562
Part-solid	1.46(0.74–2.88)	0.28	0.3770	1.72 (0.72–4.10)	0.218	0.5439
Solid	Reference		Reference	Reference		Reference
Nodule location, upper vs. middle/lower lobe	1.93 (1.14–3.27)	0.02	0.6581	2.27 (1.24–4.16)	0.008	0.8211
Nodule count, per each additional nodule <sup>a</sup>	0.92 (0.85–1.00)	0.049	−0.0824	0.99 (0.82–1.19)	0.927	−0.0142
Spiculation, yes vs. no	2.17 (1.16–4.05)	0.02	0.7729	3.40 (1.36–8.46)	0.013	1.2229
Model constant			−6.7892			−4.7364

<sup>a</sup> 62 years subtracted from actual age, 4 mm subtracted from actual nodule size and 4 subtracted from actual number of nodules in nodule count. This is done in accordance with PanCan original models to centre on PanCan mean age, nodule size and count



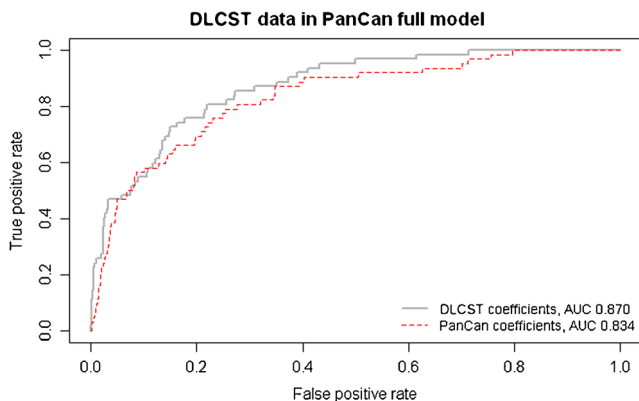
**Fig. 1** ROC curves with AUC using PanCan parsimonious model 1b on DLCST data

## Discussion

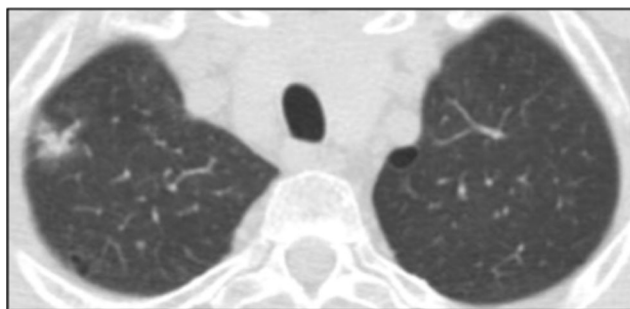
The high risk discrimination ability of the PanCan risk prediction models was largely validated by DLCST data from Denmark. However, AUCs did not reach 0.90. The specific AUCs of the different PanCan models were not reported in the original paper, and thus, exact comparisons are not possible.

Nodule size and spiculation are confirmed to be very significant risk predictor variables in both PanCan and DLCST (Tables 2, 3 and Fig. 3).

An important difference between the PanCan and the DLCST cohorts is the difference in nodule size of the included nodules; nodules with diameters of only 1–2 mm are included in the PanCan cohort, whereas 3 mm was the lower limit in DLCST. Thus, the mean diameter of benign nodules differs by almost 3 mm. We hypothesise that the stronger performance in the original PanCan risk prediction model cohort is mainly due to the inclusion of many very small nodules in the PanCan study. A nodule of 1–2 mm is hardly ever malignant, and this causes the discriminative power of the risk model to appear stronger than it actually is when it comes to cases of more doubt (bigger benign nodules).



**Fig. 2** ROC curves with AUC using PanCan full model 2b on DLCST data



**Fig. 3** Malignant solitary pulmonary nodule with spiculation and partial solid opacity. Low-dose chest screening CT from DLCST

With recent advances in automated nodule volumetry and volume doubling time measurements, a potential improvement in risk prediction by the use of these for size estimation, instead of using subjective, visual diameter measurement, could be tested.

When using DLCST data, age and family history were both significant predictors, and they should probably be added to a future parsimonious model. Age is a well-known lung cancer risk factor, and it is generally recognised that the incidence of lung cancer increases significantly with age [15]. Familial aggregation of lung cancer—across histological subtypes—has previously been documented, an effect that remains present after adjustment for confounders such as socio-economic status and smoking habits [16].

According to our logistic models, the effect of sex seems to be opposite in the original PanCan models: in the DLCST cohort, female sex appears to lower the risk of lung cancer. It is still controversial whether women have a different susceptibility to tobacco carcinogens; also it is important to consider both sex-related, i.e. biological differences, and gender-related, i.e. socially constructed, differences between men and women; the latter differ substantially between cultures and countries, and thus a generalised cross-culture risk prediction based on sex may be difficult [17]. Complex interactions of genetics, environmental and social constructions complicate conclusions [18], and these contradicting results suggest that sex should probably be removed from the model, as no true sex-related and culture-independent difference in lung cancer risk can be concluded, and it is certainly not supported by our results. The total number of malignant nodules was 102 in PanCan, 42 in BCCA and 66 in DLCST, and some nodules appeared in the same participants; therefore, as a result of the limited total numbers of included participants with lung cancer in PanCan, BCCA and DLCST, observed sex differences could be a matter of chance. Emphysema is not significantly associated with nodule malignancy in this validation cohort, nor was it in the PanCan study. It has, however, previously been shown that emphysema is significantly associated with lung cancer [19, 20], but perhaps emphysema is associated with increased risk of both benign and malignant nodules, and thus less useful in nodule-malignancy determination.

Regarding nodule count, it is possible that the different ways of counting the nodules in DLCST compared to PanCan (accumulative over the whole period of screening compared to nodule count in first scan only) is accountable for the differences in OR and *p* values observed; no effect was seen in DLCST, whereas significant reduction in lung cancer risk was seen with increasing nodule count in PanCan. Because we included all nodules found during all screening rounds in DLCST, and because the data in DLCST did not systematically include time of nodule disappearance, nodule count *per scan* could not be accurately calculated in DLCST, and thus accumulated nodule count was used instead.

In the PanCan risk prediction model paper it is stated that smoking history was not independently associated with lung cancer in the fully adjusted model and was thus left out. It is unclear how smoking history was defined (pack-years, smoking duration, etc.); also, there are other non-significant predictor variables, which indeed were included in the PanCan full model. Both tobacco exposure (pack-years) and lung function have previously been shown in several studies to be important predictors [8, 19, 20], and they should probably be included as predictor variables in future studies. Asbestos exposure and inhalation of other harmful particles could also be tested for risk predictive potential.

Lastly, we suggest further validation of the risk prediction models in cohorts with higher prevalence of malignant nodules and where small, benign nodules are less common, as we see in daily clinical practice.

## Conclusion

The PanCan risk prediction models show high lung cancer risk discrimination for solitary pulmonary nodules in the DLCST cohort, the prediction being mainly based on nodule size. However, we suggest inclusion of age and family history of lung cancer as predictor variables in the parsimonious model as well; these variables were significant predictors in DLCS T. In addition, we propose the variable sex be removed from the models, as our results did not support the PanCan conclusion that female sex is associated with increased risk; further work confirming the role of sex in risk stratification in different populations is needed.

**Acknowledgments** The scientific guarantor of this publication is Asger Dirksen. The authors of this manuscript declare no relationships with any companies whose products or services may be related to the subject matter of the article. Bram Van Ginneken works for Fraunhofer MEVIS in Bremen, Germany. This study has received funding by The Danish Ministry of Health and AstraZeneca. One of the authors has significant statistical expertise. Institutional review board approval was obtained. Written informed consent was obtained from all subjects (patients) in this study. Some study subjects have been previously reported in DLCST studies regarding lung cancer screening, lung function, lung density, airway segmentation, and visual assessment of emphysema, airway

abnormalities and interstitial abnormalities in the Danish Lung Cancer Screening Trial. However, analyses of this study are new and have not been published before. All relevant references are disclosed. Methodology: prospective, diagnostic study, performed at two institutions.

## References

1. Parkin DM, Bray F, Ferlay J, Pisani P (2005) Global cancer statistics, 2002. *CA Cancer J Clin* 2:74–108
2. Aberle DR, Adams AM, Berg CD et al (2011) Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med* 5:395–409
3. Bach PB, Mirkin JN, Oliver TK et al (2012) Benefits and harms of CT screening for lung cancer: a systematic review. *JAMA* 22:2418–2429
4. Bach PB, Kattan MW, Thornquist MD et al (2003) Variations in lung cancer risk among smokers. *J Natl Cancer Inst* 6:470–478
5. Spitz MR, Hong WK, Amos CI et al (2007) A risk model for prediction of lung cancer. *J Natl Cancer Inst* 9:715–726
6. Spitz MR, Etzel CJ, Dong Q et al (2008) An expanded risk prediction model for lung cancer. *Cancer Prev Res (Phila)* 4:250–254
7. Cassidy A, Myles JP, Liloglou T, Duffy SW, Field JK (2006) Defining high-risk individuals in a population-based molecular-epidemiological study of lung cancer. *Int J Oncol* 5:1295–1301
8. Tammemagi CM, Pinsky PF, Caporaso NE et al (2011) Lung cancer risk prediction: prostate, lung, colorectal and ovarian cancer screening trial models and validation. *J Natl Cancer Inst* 13:1058–1068
9. Swensen SJ, Silverstein MD, Ilstrup DM, Schleck CD, Edell ES (1997) The probability of malignancy in solitary pulmonary nodules. Application to small radiologically indeterminate nodules. *Arch Intern Med* 8:849–855
10. Gould MK, Ananth L, Barnett PG (2007) A clinical model to estimate the pretest probability of lung cancer in patients with solitary pulmonary nodules. *Chest* 2:383–388
11. Li Y, Wang J (2012) A mathematical model for predicting malignancy of solitary pulmonary nodules. *World J Surg* 4:830–835
12. McWilliams A, Tammemagi MC, Mayo JR et al (2013) Probability of cancer in pulmonary nodules detected on first screening CT. *N Engl J Med* 10:910–919
13. Wille MM, Thomsen LH, Dirksen A et al (2014) Emphysema progression is visually detectable in low-dose CT in continuous but not in former smokers. *Eur Radiol* 24:2692–2699
14. Pedersen JH, Ashraf H, Dirksen A et al (2009) The Danish randomized lung cancer CT screening trial—overall design and results of the prevalence round. *J Thorac Oncol* 5:608–614
15. Tas F, Ciftci R, Kilic L, Karabulut S (2013) Age is a prognostic factor affecting survival in lung cancer patients. *Oncol Lett* 5:1507–1513
16. Cote ML, Liu M, Bonassi S et al (2012) Increased risk of lung cancer in individuals with a family history of the disease: a pooled analysis from the International Lung Cancer Consortium. *Eur J Cancer* 13:1957–1968
17. Pauk N, Kubik A, Zatloukal P, Krepela E (2005) Lung cancer in women. *Lung Cancer* 1:1–9
18. Belani CP, Marts S, Schiller J, Socinski MA (2007) Women and lung cancer: epidemiology, tumor biology, and emerging trends in clinical research. *Lung Cancer* 1:15–23
19. Mizuno S, Takiguchi Y, Fujikawa A et al (2009) Chronic obstructive pulmonary disease and interstitial lung disease in patients with lung cancer. *Respirology* 3:377–383
20. Wilson DO, Weissfeld JL, Balkan A et al (2008) Association of radiographic emphysema and airflow obstruction with lung cancer. *Am J Respir Crit Care Med* 7:738–744