





# Predicting the future risk of lung cancer

Development and validation of QCancer2 (10-year risk) lung model and evaluating the performance of nine prediction models

Weiqi Liao (presenting), Carol Coupland, Judith Burchardt, David Baldwin, the DART initiative (WP6), Fergus Gleeson, Julia Hippisley-Cox



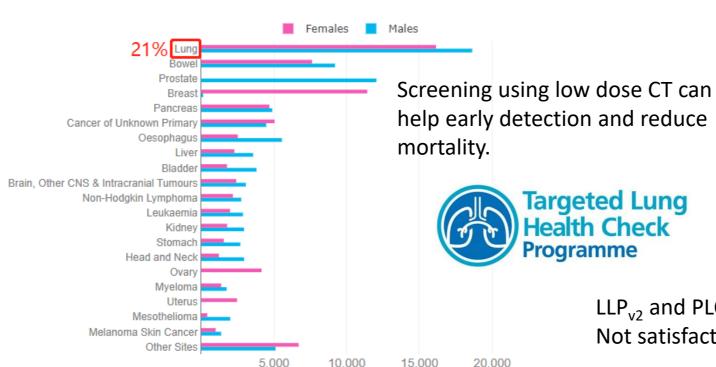






# Research background and rationale

The 20 Most Common Causes of Cancer Deaths, UK, 2017-2019



Risk prediction model
Select those at high risk for screening

Medium risk

LLP<sub>v2</sub> and PLCO<sub>m2012</sub> models in TLHC programme Not satisfactory model performance (O'Dowd et al)

Low risk

 Motivation & aim: to develop and validate a model that works better for the English primary care population for lung cancer screening.

High risk







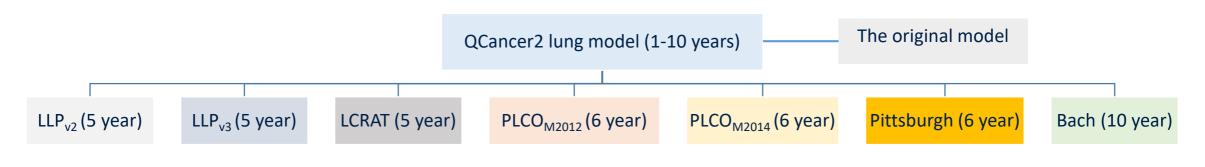
#### Methods

Model development (12.99M)

- Cox regression
- By sex (25-84 years old)

Model validation (4.14M)

- Discrimination measures: Harrell's C, D statistic, and  $R_D^2$
- Calibration plots
- Stage 1: develop and validate the QCancer2 (10-year risk) lung model
- Stage 2: model evaluation



- 1. Ever-smokers aged 55-74 years
- The population for the Targeted Lung Health Check programme

Two approaches

2. The eligibility criteria for study participants in each model

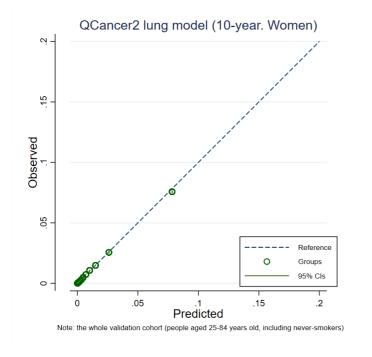


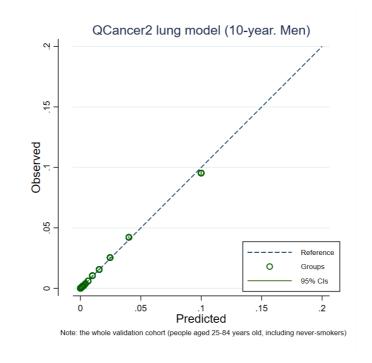




#### Results – Discrimination and calibration

Patients aged 25-84 years	Harrell's C	D statistic	$R_{\mathrm{D}}^{2}$
QCancer2 (Women)	0.897 (0.893 - 0.900)	2.81 (2.77 - 2.85)	65.4% (64.8 – 66.0)
QCancer2 (Men)	0.904 (0.901 - 0.906)	2.79 (2.76 - 2.83)	65.0% (64.5 – 65.6)





Please come to my poster (#9) to know more the predictors and the model performance when comparing with other models.







### Strengths of the QCancer2 (10-year risk) lung model

The best model performance

All ethnicities, 25-84 years and all smoking status – **Inclusive** algorithm for broad application

Robust methodology,
A representative population,
A big sample size (17M+ patients)

A **flexible** model – Allows 1-10 years predictive horizon Strengths Good face validity primary care dataset for the same population in application Allows sex-specific risk stratification A unique strength







### Conclusions & take home messages

- The QCancer2 (10-year risk) lung model has the best model performance in discrimination, calibration, and net benefit among the nine prediction models in two approaches.
- It may be more suitable to use the QCancer2 (10-year risk) lung model for lung cancer screening in the English population. (We submitted a public consultation comment to the UK National Screening Committee in early June 2022)

Email: Weiqi.Liao@phc.ox.ac.uk