

University of Groningen

Lung cancer LDCT screening and mortality reduction - evidence, pitfalls and future perspectives

Oudkerk, Matthijs; Liu, ShiYuan; Heuvelmans, Marjolein A; Walter, Joan E; Field, John K

Published in:
Nature Reviews Clinical Oncology

DOI:
[10.1038/s41571-020-00432-6](https://doi.org/10.1038/s41571-020-00432-6)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2021

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Oudkerk, M., Liu, S., Heuvelmans, M. A., Walter, J. E., & Field, J. K. (2021). Lung cancer LDCT screening and mortality reduction - evidence, pitfalls and future perspectives. *Nature Reviews Clinical Oncology*, 18(3), 135-151. <https://doi.org/10.1038/s41571-020-00432-6>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.



Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



Lung cancer LDCT screening and mortality reduction — evidence, pitfalls and future perspectives

Matthijs Oudkerk^{1,2}, ShiYuan Liu³, Marjolein A. Heuvelmans^{4,5}, Joan E. Walter⁶ and John K. Field⁷

Abstract | In the past decade, the introduction of molecularly targeted agents and immune-checkpoint inhibitors has led to improved survival outcomes for patients with advanced-stage lung cancer; however, this disease remains the leading cause of cancer-related mortality worldwide. Two large randomized controlled trials of low-dose CT (LDCT)-based lung cancer screening in high-risk populations — the US National Lung Screening Trial (NLST) and NELSON — have provided evidence of a statistically significant mortality reduction in patients. LDCT-based screening programmes for individuals at a high risk of lung cancer have already been implemented in the USA. Furthermore, implementation programmes are currently underway in the UK following the success of the UK Lung Cancer Screening (UKLS) trial, which included the Liverpool Health Lung Project, Manchester Lung Health Check, the Lung Screen Uptake Trial, the West London Lung Cancer Screening pilot and the Yorkshire Lung Screening trial. In this Review, we focus on the current evidence on LDCT-based lung cancer screening and discuss the clinical developments in high-risk populations worldwide; additionally, we address aspects such as cost-effectiveness. We present a framework to define the scope of future implementation research on lung cancer screening programmes referred to as Screening Planning and Implementation RAtionale for Lung cancer (SPIRAL).

Lung cancer is currently both the most commonly diagnosed cancer (accounting for 11.6% of all cancer diagnoses) and the leading cause of cancer-related mortality (18.4% of overall cancer mortality) in both men and women worldwide¹. In the past decade, the introduction of molecularly targeted agents and immune-checkpoint inhibitors into the therapeutic armamentarium for patients with stage IV (advanced stage) lung cancer has led to improved survival outcomes². However, these therapeutic approaches are beneficial only for restricted subsets of patients, and therefore the majority of patients with stage IV lung cancer die within 5 years of diagnosis¹. Nonetheless, patients with stage IA (early stage) disease have a >75% chance of survival over 5 years³; thus, to date, the main strategy shown to substantially reduce lung cancer mortality over a longer time period is predicated on early detection using low-dose CT (LDCT)-based screening in asymptomatic individuals^{4–7}. In settings in which lung cancer screening programmes have been implemented, annually, ~1–3% of participants are diagnosed with lung cancer, 50–70% of them with stage I disease^{7–13}. These patients usually

undergo surgery with curative intent, with other available therapeutic options being stereotactic radiotherapy, brachytherapy and percutaneous tumour ablation. Lung cancer is a tobacco-related disease: in high-income countries, ~10–20% of current and former heavy smokers will be diagnosed with lung cancer during their lifetime compared with 1–2% of never smokers^{14,15}. Thus, individuals with a history of smoking are likely to derive the greatest benefit from screening.

In this Review, we discuss the current evidence supporting the effectiveness of implementing national lung cancer screening programmes as well as how such programmes should be developed in the future. We focus on aspects such as the identification of the target population, participant recruitment and compliance, screening frequency, integrated smoking cessation interventions, cost-effectiveness and sex differences. We also present an overview of current lung cancer screening programmes worldwide and discuss the future opportunities to leverage artificial intelligence (AI) in LDCT-based lung cancer screening. All of these areas should be considered in the scope of future implementation research programmes

 e-mail: m.oudkerk@rug.nl
<https://doi.org/10.1038/s41571-020-00432-6>

Key points

- The role of CT lung cancer screening on lung cancer-related mortality reduction has been under debate for five decades and is now an evidence-based reality.
- The implementation of low-dose CT (LDCT)-based screening requires an optimal risk modelling methodology in order to select the high-risk population that will derive the greatest benefits.
- Biomarkers have the potential to be included in future risk assessment models and work-up of CT nodules; laboratory tests should be developed to improve risk assessment before CT screening.
- LDCT-based lung nodule diameter measurement cannot be used as an imaging biomarker for effective lung cancer risk stratification; however, volume doubling time can effectively be used as an imaging biomarker to rule out lesions with benign tissue growth.
- The LDCT lung cancer screening interval should eventually be tailored to the expected mean nodule growth of the targeted population, starting with personalized screening, to improve screening performance (in particular for women).
- Effective smoking cessation interventions must be integrated within cost-effective lung cancer screening programmes.
- We propose a framework — the Screening Planning and Implementation Rationale for Lung cancer (SPIRAL) — to define the scope of future implementation research on lung cancer screening programmes.

on lung cancer screening, in a framework we refer to as the Screening Planning and Implementation Rationale for Lung cancer (SPIRAL) (FIG. 1).

Identification of a high-risk population

To minimize the potential harms associated with cancer screening (such as exposure to radiation) and maximize its effectiveness, screening programmes should be limited to individuals who are at high risk of a particular cancer within the general population¹⁶. Typically, screening programmes are focused on a prespecified subset of individuals within the general population on the basis of either age (for example, in colorectal cancer screening) or a combination of age and sex (such as in breast and cervical cancer screening). For lung cancer screening, an improved approach has been implemented in the UK through several lung cancer CT screening implementation studies: the Liverpool Healthy Lung Project¹⁷, Manchester Health Lung Check¹⁸, the West London Cancer Screening pilot¹⁹ and the Yorkshire Lung Screening trial²⁰. Besides age (>60 years), smoking status has been shown to have the greatest influence on the probability of developing lung cancer (with odds ratios of 2.17 (1.21–3.85) and 15.25 (5.71–40.65) for smoking durations of 1–19 years and ≥60 years, respectively)²¹. However, several other factors also contribute to this risk,

including a family history of lung cancer (especially for individuals aged <60 years, who have an odds ratio of 2.02 (1.18–3.45) compared with 1.18 (0.79–1.76) for individuals with a family history and aged ≥60 years)²¹ and an individual history of other respiratory diseases, other malignancies and exposure to asbestos. Other risk factors have been reported in the literature, including exposure to radon and a number of other carcinogens (such as diesel exhaust fumes); however, to date, they have not been included in any of the validated risk models.

The use of prediction models integrating several risk factors in lung cancer-screening research has gained credence over the past 10 years. Indeed, the use of validated risk models is integral to all current screening and early detection programmes in Europe. Several multivariable risk-prediction models have been published and reviewed²²; however, only two — PLCO_{M2012} (REF.²³) and LLP_{v2} (REF.²⁴) — have thus far been used to guide the selection of participants in lung cancer screening clinical trials and projects. The US National Lung Screening Trial (NLST) involved current and former heavy smokers (≥30 or more pack-years of cigarette smoking history; former smokers were involved if they quit smoking <15 years before) aged 55–74 years. These participants were randomly allocated to undergo three annual rounds of screening with chest LDCT or single-view chest radiography⁴. The NLST dataset has been analysed using several risk-prediction models, leading to the conclusion that the NLST selection criteria and the United States Preventive Services Taskforce (USPSTF) criteria recommendations for lung cancer screening could have been greatly improved if a risk model incorporating variables beyond age and smoking history had been implemented^{25–27}.

Currently, LLP_{v2} is the only risk model that has been used to select participants in a randomized controlled trial (RCT) of lung cancer screening: in the UK Lung Cancer Screening (UKLS) trial²⁸, a 5-year lung cancer risk of ≥5% according to LLP_{v2} was used as an inclusion criterion, together with an age of 50–75 years. Participants in this trial underwent LDCT-based screening or no screening. The percentage of participants with lung cancer identified in the LDCT arm at baseline (1.7%) was higher in the UKLS trial than in the NLST or NELSON (1.03% and 0.9%, respectively)²⁸.

Of note, NELSON involved current and former heavy smokers (≥30 pack-years) aged 55–75 years who were randomly allocated to several rounds of chest LDCT-based screening or no screening. The LLP_{v2}-based criteria used in the UKLS trial were subsequently adopted to select participants in the Liverpool Healthy Lung Programme¹⁷. The UK is currently leading the way in Europe in terms of implementing lung cancer early detection with LDCT-based screening, with major programmes ongoing in the Liverpool¹⁷, Manchester¹⁸, Yorkshire²⁰ and London^{19,29} regions. Moreover, in 2019, NHS England provided major investment to introduce a national programme in 10 new regions³⁰. These new programmes will involve a combination of both the PLCO_{M2012} and LLP_{v2} risk models to recruit participants, thus demonstrating interest in targeted recruitment approaches.

Author addresses

¹Institute of Diagnostic Accuracy (iDNA), Groningen, Netherlands.

²Faculty of Medical Sciences, University of Groningen, Groningen, Netherlands.

³Department of Radiology, Changzheng Hospital, Second Military Medical University, Shanghai, China.

⁴Department of Pulmonary Diseases, Medisch Spectrum Twente, Enschede, Netherlands.

⁵Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, Netherlands.

⁶Institute of Diagnostic and Interventional Radiology, University Hospital Zurich, Zurich, Switzerland.

⁷Roy Castle Lung Cancer Research Programme, Department of Molecular and Clinical Cancer Medicine, University of Liverpool, Liverpool, UK.

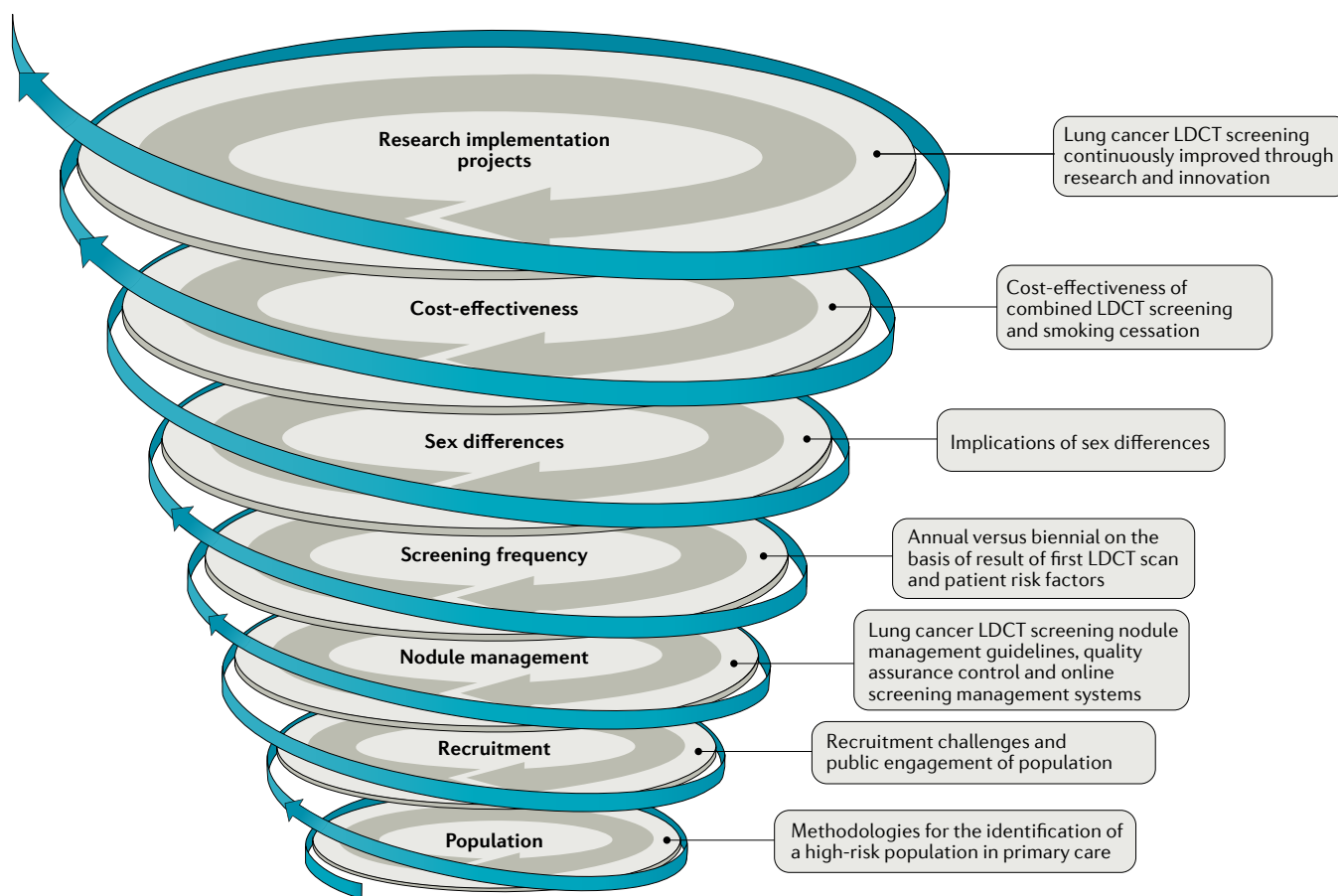


Fig. 1 | **Screening Planning and Implementation RAionale for Lung cancer.** Herein, we present a framework to define the scope of future implementation research on lung cancer screening programmes, referred to as Screening Planning and Implementation RAionale for Lung cancer (SPIRAL). LDCT, low-dose computer tomography.

In one of the most comprehensive analyses, nine different risk models were used to analyse data from the Prostate, Lung, Colon, and Ovarian (PLCO) cancer screening trial and NLST datasets²². The selected sophisticated models incorporated well-documented risk variables (such as family history of lung cancer, previous malignancy, previous respiratory disease and exposure to asbestos). However, not all risk factors were considered in these comparisons, which were only based on age, sex and tobacco-related factors, thus underestimating the lung cancer risk of never smokers. The PLCO_{M2012} model had the best predictive performance in this analysis, with an area under the curve (AUC) of >0.77. Several studies have also shown the cost-effectiveness of screening in high-risk populations, leading to the conclusion that improved risk-prediction models would further reduce costs per life years (LYs) saved^{22,31}. The cost-effectiveness analysis only revealed a modest gain of additional LYs. In addition, the use of lung cancer prediction models increased the risk of overdiagnosis owing to the preferential selection of older individuals; thus, the researchers concluded that the future development of risk-based lung cancer screening needs to incorporate life expectancy³¹.

Using risk models in national screening programmes has potential limitations that must be acknowledged.

In particular, information on risk variables has to be either available in primary health-care records or obtained directly from the patient. The collection of these data in the UK implementation studies has involved a two-step process, whereby all patients with a smoking history in the primary care notes and/or electronic health records were invited and were then provided with a structured questionnaire following their consent to be included in the studies. However, smoking history is not always recorded in primary care notes and might thus be challenging in other countries and not a feasible approach. The advent of social media and the use of clinical apps might provide solutions to obtain information on risk variables directly from patients; however, these approaches remain in the early stages of development³².

Currently, none of the validated prediction models to identify individuals with a high risk of lung cancer has incorporated biomarkers or susceptibility genes, even though major efforts have been undertaken in this regard³³. Integral, a major lung cancer programme from the NIH³⁴, is currently focused on this topic and has generated some early encouraging data on the integration of genetic susceptibility pathways^{35–37} and circulating biomarkers³⁸ in risk-prediction models. Indeed, the next stage in the development of risk-prediction models will have to move beyond epidemiological and clinical data

to also include validated biomarkers. This active area of research will require access to current CT-screening biobanks as well as the development of high-quality prospective biobanks embedded in future screening programmes together with radiomics data (volume and density growth characteristics). Future molecular tests not only need to be validated but must also be cost-effective, possibly using nanotechnology-based approaches³⁹.

Recruitment and adherence issues

The real-world experience in the USA, where only a fraction (<5%) of individuals at high risk of lung cancer are screened, demonstrates the difficulties in the effective recruitment of participants in national screening programmes even when they are endorsed by most major medical societies⁴⁰. The challenges of recruitment and screening adherence differ between regions because they depend on the nature of the health-care system as well as on the public and physician opinions on screening — clearly, a unique approach has to be chosen for each country. Nevertheless, two principles should be common to all approaches to recruitment: screening should only be implemented for high-risk individuals, and the appropriate presentation of potential benefits and risks is crucial⁴¹. Experience from the UKLS trial has revealed that, especially in the first stage of recruitment, current smokers and individuals from lower socio-economic groups are least inclined to participate^{42,43}. For current smokers, emotional barriers seem to represent a central obstacle to screening participation⁴². More than ever, primary care physicians could be the focal point in ensuring screening uptake by individuals who are mostly likely benefit⁴⁰. Other major contributors to the low uptake of screening might be the false-positive rate (that is, of detection of a non-malignant nodule; 24%¹⁶) that was reported in the NLST⁴ as well as the perceptions of some patients and carers⁴⁴. However, in the NELSON trial, with results published 9 years after those from the NLST⁴⁵ and incorporating optimized nodule-management protocols and risk-stratification algorithms, the false-positive rate was only 1.2% and the referral rate only 2.1%¹⁰. Of note, the definition of a positive screen result differed between the NLST and NELSON study: in the NLST, the two possible outcomes of a chest LDCT or radiography were ‘negative’ and ‘positive’ whereas, in NELSON, ‘indeterminate’ was introduced as a new outcome⁴⁶. Only when indeterminate nodules were found to have grown at a short-term follow-up LDCT scan was the indeterminate screen result reclassified as positive.

Eventually, the successful recruitment of individuals at high risk of lung cancer will depend on the combined efforts of primary-care physicians and specialists. In order to ease the pressure on the former, the responsibility for determining an individual’s eligibility has to be considered as a multidisciplinary activity and, thus, discussions around shared decision-making, counselling for smoking cessation and potential treatment options should be combined across clinical specialties.

Challenges in the recruitment of high-risk and hard-to-reach individuals remain one of the major

barriers to the implementation of lung cancer screening programmes⁴⁷. Even among the most efficient centres in terms of recruitment in ongoing UK implementation projects, few have a participation rate of >50%^{17–19}.

Radiological evidence

Before the 2010s, the technical performance of chest radiography, alone or in combination with sputum cytology, was evaluated in population-based lung cancer screening programmes^{16,41}. However, these studies did not show reductions in lung cancer mortality, and the screening method was proven not to be sensitive enough^{16,48,49}. In the 2000s, the introduction of LDCT renewed interest in assessing the performance of imaging-based lung cancer screening approaches¹⁶. A chest LDCT entails a radiation dose of ~1.5 mSv, which is 15-fold higher than the dose delivered to obtain a conventional chest X-ray but <25% of that delivered with conventional chest CT⁵⁰.

While other diagnostic methods, such as MRI or genetic testing, have been explored in population lung cancer screening, results from RCTs that would support their use in current clinical practice are not available^{16,51,52}. Currently, LDCT-based lung cancer screening is the only screening approach that has resulted in a statistically significant reduction of lung cancer-related mortality in two independent, sufficiently powered RCTs^{4,10} (FIG. 2). In 2011, researchers from the NLST reported a 20.0% reduction in lung cancer-related mortality after a median follow-up of 6.5 years ($P=0.004$) in patients undergoing three annual LDCT-based screenings compared with those undergoing chest radiography screening with the same frequency⁴. The overall mortality reduction in the LDCT group was 6.7% ($P=0.02$). In 2020, results from the NELSON trial showed a cumulative rate ratio (RR) for death from lung cancer in men of 0.76 (95% CI 0.61–0.94) in the screening arm relative to the control arm at 10 years¹⁰. The cumulative RR for all-cause mortality was 1.01 (95% CI 0.92–1.11). Nevertheless, the implementation of LDCT in screening programmes is still ongoing in the USA and anticipated in Europe in the next decade^{16,41,53,54}.

Nodule prevalence and risk stratification. Effective risk stratification and management of detected lung nodules is crucial for the success of any lung cancer screening programme. Baseline nodules with an unknown developmental timeframe need to be distinguished from new nodules (after baseline) that have developed within a known timeframe⁵⁵. Depending on the detection limit, 22–51% of participants in screening RCTs have a lung nodule detected at baseline^{11,56–64}. Furthermore, available data from the Early Lung Cancer Action Project (ELCAP)⁶⁵, the International (I)-ELCAP⁶², the Pittsburgh Lung Screening Study⁵⁶, the Mayo trial⁶⁶, the NLST⁶⁷ and the NELSON trial⁵⁵ suggest that, annually, 3–13% of participants develop a new nodule (negative or positive) after any baseline screening. Importantly, most lung nodules detected, either at baseline or thereafter, are small. Data from lung cancer screening trials with none or a very low detection limit (>3 mm or >15 mm³; Mayo trial⁶⁶, ELCAP⁶⁵, I-ELCAP⁶² and NELSON⁵⁵)

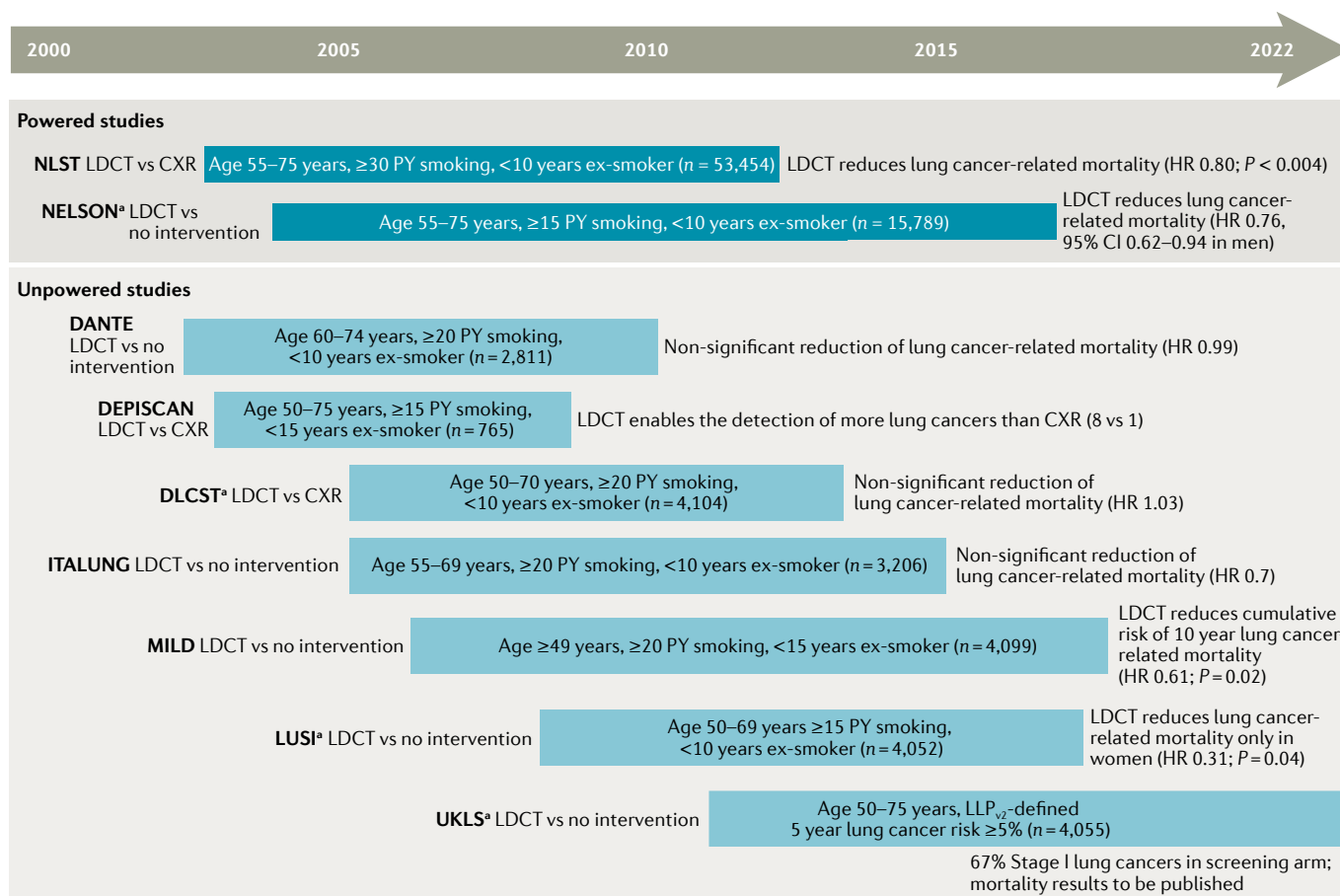


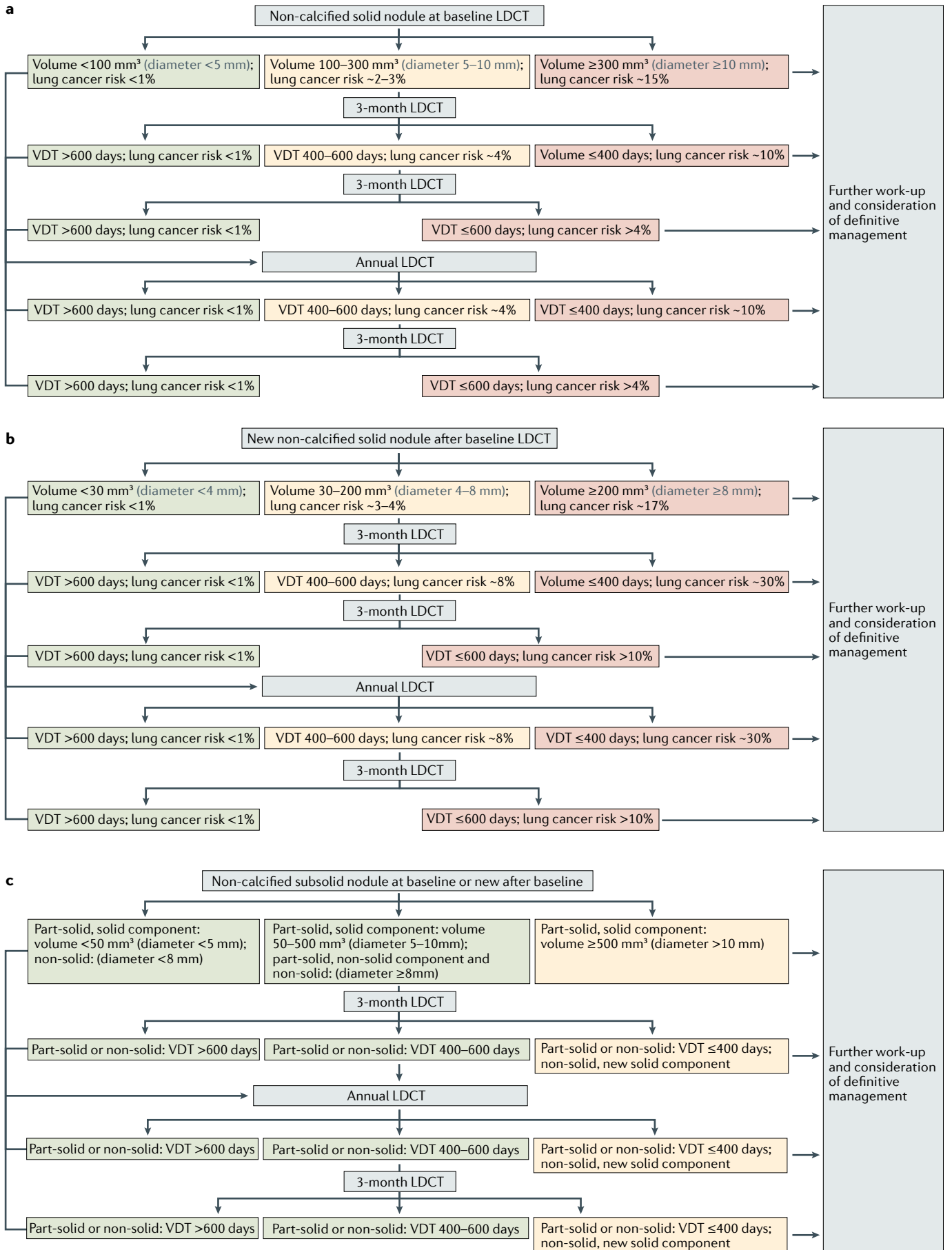
Fig. 2 | **Randomized controlled trials of LDCT-based approaches to lung cancer screening.** Timeline of randomized-controlled trials (RCTs) of low-dose computer tomography (LDCT)-based lung cancer screening, showing the time from the recruitment date to the end of follow-up and relevant findings associated with each trial. Of note, the mortality data are expressed with a P value for all RCTs except NELSON, for which confidence levels are provided. CXR, chest radiography; PY, packet years. ^aRCTs performing CT volumetry.

suggest that >50% of the detectable lung nodules have a volume of <50 mm³ or a maximum diameter of <5 mm (REFS^{55,57–60,65,66,68}). Similarly, NLST (with a detection limit of 4 mm for the longest diameter) revealed a baseline nodule prevalence of 51% for nodules of 4–6 mm. The detection of multiple nodules is common in screening practice: ~50% of participants have more than one nodule at baseline, and >20% of those who develop new nodules have multiple nodules; each nodule requires a separate risk assessment^{69,70}.

Nodule size assessment. An accurate and reproducible assessment of nodule size is central to ensuring appropriate nodule management. The assessment of nodule size has been routinely based on the manual measurement of the longest diameter^{71,72}. Nevertheless, this approach was shown to be unreliable when compared with subsequent methods, such as volumetry, because pulmonary nodules are seldom perfectly geometrically shaped^{73,74}. Several European lung cancer screening trials (FIG. 2) have incorporated volumetry involving semi-automated volume estimation after 3D reconstruction of thin CT slices of nodules^{6,10,51,55,60}. This approach was advocated in the European Statement on Lung cancer Screening

(EUPS)¹⁶ and was subsequently implemented in clinical practice guidelines from the British Thoracic Society, suggesting that, whenever available, volumetry should be preferred to diameter measurements^{16,71,75,76}. Moreover, in the 2019 Lung Imaging Reporting and Data System (Lung-RADS) screening guidelines, volume standards have been added as a more reproducible alternative to manual linear measurements when appropriate software is available^{16,53}.

Nodule growth. With the use of appropriate size cut-offs, most nodules detected during lung cancer screening can be classified as low-risk or intermediate-risk nodules, and decisions can be made on additional follow-up screens (regular (1 year) or short-term (3 months) according to the EUPS¹⁶). At follow-up screens, risk stratification should be based on nodule growth^{16,71,75,76}. Again, considering that most nodules detected in lung cancer screening are very small, tumour growth assessment based on 2D diameter evaluation has been considered unreliable compared with volumetry⁷⁶. For example, in the Lung-RADS screening guidelines, growth has been defined as an increase of >1.5 mm in diameter or of >2 mm³ in volume⁵³. In a spherical nodule with a diameter



◀ Fig. 3 | **The NELSON-Plus Protocol for LDCT scan-detected lung nodules.** **a** | Non-calcified solid nodules detected at baseline low-dose computer tomography (LDCT) scans. **b** | New non-calcified solid nodules detected after baseline LDCT scans. **c** | Non-calcified subsolid nodules detected at baseline or new nodules detected after the baseline scan. These risk-stratification protocols are based on the data from the NELSON trial^{16,55,70,80,83}. Diameters do not correspond to the volumes, but to former reference values. VDT, volume-doubling time.

of 5 mm (and thus, a volume of ~65 mm³), a diameter increase to 6.5 mm would result in a more than doubled volume (144 mm³), whereas a volume increase to 67 mm³ corresponds to a diameter increase to only 5.04 mm. An analysis of 2,240 intermediate-size nodules (defined as 50–500 mm³ in volume or ~4.5–10 mm in longest diameter), revealed a median intranodular diameter variation of 2.8 mm, above the 1.5 mm growth threshold, when volume was estimated based on the maximum versus the minimum diameter⁷⁴. Even when nodule diameter was measured semi-automatically, the intranodular variation was ≥2 mm in 85% of nodules^{74,76,77}. Importantly, volume measurements have a significantly worse performance in areas with ground glass opacity and in the measurement of part-solid nodules⁷⁸. In Asian populations, in which such nodules are more common⁷⁹, volume measurements alone might therefore not be the best option in nodule management; rather, a combination of the measurement of volume, mass and diameter of these subsolid nodules might be preferable. Another advantage of volumetry is that it enables the calculation of the volume-doubling time, a widely used surrogate for growth speed¹⁶, as opposed to considering a fixed size increase, which translates into different growth speeds at different nodule sizes. Even compared with software-guided and optimized diameter measurements, a protocol based on semi-automated nodule volume and volume-doubling time measurements yielded the highest specificity (94.9% versus 90.0% with the diameter-based protocol) and a positive predictive value (14.4% versus 7.9% with the diameter-based protocol), with a similar negative predictive value (99.9% in both protocols), in an analysis of data from the NELSON trial^{76,80}.

Nodule subtypes. Radiological detection enables the classification of pulmonary nodules into non-calcified pulmonary nodules, which comprise solid and subsolid nodules, the latter including ground-glass (non-solid) nodules and part-solid nodules, and calcified nodules. From the perspective of lung cancer screening, this distinction is relevant for two reasons. Firstly, both at baseline screening and in subsequent rounds of screening, (new) subsolid nodules are considerably less prevalent than (new) solid nodules and, overall, <10% of lung cancer screening participants present with non-solid nodules^{16,81–83}. Secondly, compared with solid nodules, non-solid nodules (including pre-malignancies) are associated with an equivalent or a higher prevalence of lung cancer but their indolent nature (they are nearly always stage I cancers or in the pre-stages of lung cancer) has been shown both in prospective studies and RCTs^{79,81,82,84–87}. Data from the NELSON trial and EUPS have formed the basis to develop risk-stratification protocols for different LDCT screen-detected nodules¹⁶ (FIG. 3).

Overdiagnosis and false positives. The identification of clinically significant lung cancer while preventing overdiagnosis and false-positive results is a central challenge in LDCT-based lung cancer screening. In this regard, clinical decision-making upon the detection of subsolid nodules is particularly challenging because they are more often malignant than solid nodules but have a slower growth rate⁷⁹. Therefore, continuous benchmarking of risk-stratification algorithms is essential. For example, a comparison of the screening results from NELSON (using a volume-based protocol) and NLST (using a diameter-based protocol) showed substantial differences in false-positive baseline screening results, with positive-test rates of 2.1% versus 24%, positive predictive values of 43.5% versus 3.8%, and false-positive rates of 1.4% versus 23.3%^{76,77}.

Nodule-based risk-prediction models

The potential of integrating nodule data from LDCT scans with the patient's clinical and epidemiological information has enabled the development of nodule-based lung cancer risk models. In certain instances, these models have been used in clinical practice not only to manage the radiological diagnostic follow-up but also to calculate the most appropriate time for a follow-up scan.

In 2013, researchers from the Pan-Canadian Early Detection of Lung Cancer Study (PanCan) published several nodule-based risk-prediction models⁸⁸, now referred to as the Brock parsimonious model (PanCan-1b) and the comprehensive model (PanCan-2). These models were developed using data from 1,871 participants in PanCan and validated using a dataset comprising 1,090 individuals involved in chemoprevention trials from the British Columbia Cancer Agency. These two high-risk screening cohorts had been followed up for a minimum of 2 years to determine the probability of pulmonary nodules detected in LDCT screens being cancerous. A cancer diagnosis was associated with female sex ($P \leq 0.02$), larger size of the nodule ($P < 0.001$), location of the nodule in the upper lung ($P \leq 0.02$) and nodule spiculation ($P \leq 0.02$). These researchers also developed so-called 'full models', which additionally included older age, a family history of lung cancer, emphysema, lower nodule count and part-solid nodules as compared with solid nodules. These models had very good predictive accuracy, with AUCs of >0.94 in the external validation cohort, and thus became the management tool recommended by EUPS and the British Thoracic Society^{16,75}. In 2019, PanCan included nodule volume in these models⁸⁹. Both the diameter-based and volume-based models showed very good overall predictive performance in the test and validation datasets, with an accuracy similar to that of the previously validated PanCan models: the computer-aided detection (CAD)-assessed mean diameter and volume models both had median AUCs of 0.947 in the PanCan data and of 0.810 and 0.821, respectively, using the NLST dataset⁸⁹.

The UKLS dataset has also been used to develop a parsimonious model to estimate the probability of malignancy in lung nodules detected at baseline and at 3-month and 12-month repeat screens⁹⁰. The covariates

found to enable the prediction of lung cancer included female sex, asthma, bronchitis, asbestos exposure, a history of cancer, a family history of early and late onset of lung cancer, smoking duration, lung forced vital capacity, and nodule type (pure ground-glass and part-solid) and larger volume (measured by semi-automated volumetry). The final model incorporating all predictors had excellent discriminatory value (with an AUC of 0.885). Internal validation suggested that the model would discriminate well when applied to new data in the future (with an AUC of 0.882) and had good calibration when used with 'bootstrapping' optimization techniques.

A number of groups have attempted to develop other nodule-based risk-prediction models. The Pittsburgh Lung Screening Study cohort constitutes one such approach using probabilistic graphical models to integrate demographics, clinical data and LDCT scan-related features⁹¹. The investigators noted that the number of nodules and blood vessels as well as the number of years since the individual quit smoking were sufficient to discriminate malignant from benign nodules, with statistically significant coefficients ($P < 0.05$). The incorporation of LDCT scan-related features greatly enhanced the predictive accuracy of this model, improving cancer detection over existing methods, in particular the Brock parsimonious model ($P < 0.001$). The most notable observation of this study is that the incorporation of information on the number of surrounding vessels significantly improves on the predictive efficiency of previous models⁹¹.

In the German Lung Cancer Screening Intervention (LUSI) trial⁶, 4,052 long-term smokers aged 50–69 years were randomly allocated to undergo five annual rounds of LDCT-based screening or no screening. Data from this trial were used with the aim of validating several nodule-based risk models, such as the PanCan⁸⁹, Mayo Clinic⁹², Peking⁹³ and UKLS⁹⁰ models, using sophisticated statistical tools⁹⁴. PanCan-1b was found to be the model with the most predictive value in this validation exercise (AUC 0.93), and the UKLS model was considered the least optimal (AUC 0.58), although the study design did not consider that some of the UKLS model parameters were not available in the original LUSI dataset such as family history and exposure to asbestos. Leaving these two variables out and keeping the other coefficients in the model unchanged would most likely result in biased estimates. The editorial associated with this publication outlined the benefits and disadvantages of attempting validation of these risk models⁹⁵. This study exemplifies the importance of including parameters with a low risk of inter-reader variability in risk models. The inclusion of parameters with a high risk of inter-reader variability, such as the diagnosis of bronchitis or discrimination between part-solid and non-solid lung nodules, might strongly reduce the performance of these models for predicting outcomes in cohorts others than those with which they were developed⁹⁵. Of note, all the models discussed herein have a reduced performance when used on nodules newly detected after baseline, confirming the need for separate management protocols for these nodules.

Screening frequency

In LDCT-based lung cancer screening, the duration of the interval between two regular screening rounds (referred to as screening interval) is a crucial determinant of the benefit-to-harm ratio. By prolonging this interval, cumulative radiation and diagnostic costs decrease but the probability of a cancer diagnosis outside of the screening programme (so-called 'interval cancers') and/or that of detecting late-stage lung cancer increase. In the USA, lung cancer screening is currently performed through annual LDCTs in high-risk patients according to the USPSTF recommendations, which in turn are based on the NLST criteria⁹⁶. All countries recommend an annual screening interval; however, the outcomes of the NELSON study suggest that a sex-specific interval could be applied in the future because nodules tend to have a slower growth rate in women than in men¹⁰.

With the increasing interest in patient-tailored medicine, the question has arisen of whether decisions regarding future screening rounds should be made based on the baseline screening result, enabling the identification of subgroups of patients with lower lung cancer risk who might benefit from a biennial screening interval. To date, evidence from three screening trials (NLST, the Multicentric Italian Lung Detection (MILD) trial and NELSON) contributes to this debate.

Patz et al.⁹⁷ retrospectively evaluated the value of annual follow-up LDCT after a negative baseline screening result among participants in the NLST. Among 19,066 NLST participants with a negative baseline result, 441 (2%) were eventually diagnosed with lung cancer. Lung cancer was diagnosed within 2 years after the baseline scan in 92 individuals (0.48%, with 30 interval cancers and 62 screen-detected cancers), 52% with stage I–II cancers. An additional 118 participants had lung cancer diagnosed ≤ 1 year after the third screening (thus 2–3 years after the baseline scan), with 60% having stage I–II cancers. Owing to the very low incidence of lung cancer in the first annual screening round after baseline in participants with a negative baseline LDCT scan, Patz et al.⁹⁷ concluded that annual screening might be superfluous in these situations.

In the MILD trial, participants with a high risk of lung cancer (49–75 years of age, who smoked ≥ 20 pack-years, and were current smokers or quit < 10 years before recruitment) were randomly assigned to undergo annual screening ($n = 1,190$), biennial screening ($n = 1,186$) (both for a median follow-up of 6 years) or no screening (control group, $n = 1,723$). Ten years after the baseline scan, LDCT-based screening (annual and biennial combined) was associated with a significant reduction (39%) in lung cancer-related mortality (HR 0.61, 95% CI 0.39–0.95; $P = 0.017$) as well as a non-significant decrease (20%) in all-cause mortality (HR 0.80, 95% CI 0.62–1.03; $P = 0.069$)⁷. In an additional analysis of the results of MILD, Pastorino et al.⁹⁸ showed that lung cancer-related mortality (HR 1.10, 95% CI 0.59–2.05) and overall mortality (HR 0.80, 95% CI 0.57–1.12) at 10 years after baseline screening were similar for participants in the biennial and annual LDCT arms. In this trial, the biennial screening protocol enabled the avoidance of up to 44% of follow-up LDCT scans, without

an increase in the occurrence of stage II–IV or interval lung cancers. Although the sample size of the MILD trial was underpowered, these results suggest that individuals with a negative baseline result might benefit from undergoing biennial instead of annual screening.

Another approach to addressing whether the screening interval should be considered on an individual basis has been proposed based on a logistic regression model of lung cancer risk at the second annual screen or in the following year. This model, which was only tested retrospectively, included participants' characteristics and radiological observations, such as nodule characteristics at the first screen, using NLST data⁹⁹. For different risk thresholds, Schreuder et al.⁹⁹ projected that 2,558 (10.4%), 7,544 (30.7%), 10,947 (44.6%), 16,710 (68.1%) and 20,023 (81.6%) of 24,368 second screens could have been omitted, at the cost of delaying the diagnosis of 0 (0.0%), 8 (4.6%), 17 (9.8%), 44 (25.3%) and 70 (40.2%) of 174 lung cancers, respectively, thus concluding that the screening interval could be extended for certain participants.

In NELSON, the effect of prolonged screening intervals was studied by incorporating different intervals between each repeat round of screening. Participants randomly allocated to the LDCT arm were screened at baseline (year 1) and then in years 2, 4 and 6.5, resulting in one annual screening round, one biennial screening round and one 2.5-year screening round, respectively. The probability of lung cancer 2 years after the baseline scan was determined. Participants with a negative baseline CT, with a newly proposed cut-off volume for the largest nodule <100 mm³, had a similar, very low risk of being diagnosed with lung cancer within 2 years as participants without any baseline nodule (0.6% versus 0.4%, respectively). For participants without any baseline nodule, the 2-year probability of lung cancer was significantly lower than that of participants with intermediate-risk nodules (100–300 mm³; 2.4% probability) or high-risk nodules (>300 mm³; 16.9%) at baseline,

again suggesting that an annual LDCT after a negative baseline CT might not be necessary for some patients. In the NELSON study, the number of interval cancers and stage II–IV lung cancers detected after a screening interval of 2.5 years was higher than those detected at the annual and biennial screening rounds (FIG. 2), indicating that a screening interval of >24 months might be too long¹⁰⁰.

These studies show the added value of patient stratification based on the results from the baseline LDCT scan; however, this stratification approach leads to questioning of the value of risk assessment before testing. For example, if an individual is already eligible for screening, is further stratification according to the baseline screen a correct approach? If the baseline result was negative, should this patient not have been invited for screening at all? Nevertheless, using the baseline screening result as an additional, independent, lung cancer risk stratification together with variables specific for each participant to determine eligibility might help to reduce the number of unnecessary screenings¹⁰¹ (TABLE 1).

Sex differences in lung cancer

Lung cancer screening trials have revealed differences in lung cancer-specific mortality between men and women. Shortly after the publication of the NLST mortality results at a median follow-up duration of 7.5 years, a detailed analysis of these results stratified by several factors was presented¹⁰². In this analysis, lung cancer incidence and mortality was evaluated up to 31 December 2009 instead of 15 January 2009, the date in the original publication⁴. Lung cancer screening was found to be beneficial to a higher extent in women than in men (TABLE 2), although this interaction was not statistically significant (*P* = 0.08). Updated results from NLST were published after an extended follow-up duration of 11.3 years for lung cancer incidence and of 12.3 years for mortality¹⁰³, a period in which potential confounding owing to participation in the screening programme would be diluted. The investigators found a beneficial effect for women, with lung cancer mortality RRs in dilution-adjusted analysis of 0.89 (95% CI 0.80–1.00), 0.95 (95% CI 0.83–1.10) and 0.80 (95% CI 0.66–0.96) in the overall study population, men and women, respectively; however, when directly compared, the difference between men and women was not statistically significant. The number of patients with stage IV disease was 468 in the LDCT arm versus 597 in the radiography arm; this difference was larger for women (165 and 232 patients with stage IV disease in the LDCT and radiography arms, respectively) than for men.

The NELSON outcomes published after 10 years of follow-up were focused on the effect of screening in men owing to the low number of women involved in the trial (TABLE 2). Nevertheless, lung cancer-specific mortality outcomes were more favourable for women than for men, although the 10-year lung cancer-specific mortality results were not statistically significant in women (RR 0.67, 95% CI 0.38–1.14). At 7, 8 and 9 years after baseline LDCT-based screening, the magnitude of lung cancer-specific mortality reduction was greater

Table 1 | Follow-up recommendations based on lung cancer nodule risk

Risk category ^a	Lesions detected ^b	Recommendation
Intermediate to high risk of lung cancer (<1%)	No baseline nodule; no new nodule at follow-up screening; solid baseline nodule <100 mm ³ or <5 mm; new solid nodule <30 mm ³ or <4 mm	Consider prolonged screening interval of up to 24 months
High risk of lung cancer (~3%)	Solid baseline nodule 100–300 mm ³ or 5–10 mm; new solid nodule 30–200 mm ³ or 4–8 mm; growing solid nodule with VDT of 400–600 days; subsolid nodule, baseline or new, of any size ^c	Short-term follow-up (3 months); if negative: annual screening
Very high risk of lung cancer (>15%)	Solid baseline nodule >300 mm ³ or >10 mm; new solid nodule >200 mm ³ or >8 mm; growing solid nodule with VDT <400 days; subsolid nodule that is growing or has an altered morphology	Referral to MDT for work-up; if negative: annual screening

MDT, multidisciplinary team; VDT, volume-doubling time. ^a2-year probability of lung cancer, based on largest or fastest-growing nodule. ^bSize expressed as volume (mm³) or longest diameter (mm). ^cIn case of negative follow-up CT scan (no growth), consider prolonged screening interval up to 24 months. Table used with permission of AME publishing, from Heuvelmans, M. A. & Oudkerk, M. Appropriate screening intervals in low-dose CT lung cancer screening. *Transl Lung Cancer Res.* 7, 281–287 (2018); permission conveyed through Copyright Clearance Center, Inc.

Table 2 | Results of randomized controlled trial of lung cancer screening stratified by sex

Study	Median follow-up	Outcomes in men	Outcomes in women
NLST ¹⁰²	7.5 years after baseline scan	Screening arm: 311 lung cancer-related deaths among 15,769 individuals (2.0%); no-screening arm: 337 lung cancer-related deaths among 15,761 individuals (2.1%); RR 0.92 (95% CI 0.8–1.08) ^a	Screening arm: 158 lung cancer-related deaths among 10,953 individuals (1.4%); no-screening arm: 215 lung cancer-related deaths among 10,969 individuals (2.0%); RR 0.73 (95% CI 0.6–0.9) ^a
NLST ¹⁰³	11.3 years after baseline (incidence), 12.3 years after baseline (mortality)	Screening arm: 733 lung cancer-related deaths among 15,769 individuals (4.6%); no-screening arm: 755 lung cancer-related deaths among 15,761 individuals (4.8%); RR 0.97 (95% CI 0.87–1.07) ^a	Screening arm: 414 lung cancer-related deaths among 10,953 individuals (3.8%); no-screening arm: 481 lung cancer-related deaths among 10,969 individuals (4.4%); RR 0.86 (95% CI 0.75–0.98) ^a
NELSON ¹⁰	10 years after baseline	Screening arm: 341 lung cancers detected in 6,583 individuals (5.2%), 156 lung cancer-related deaths in 6,583 individuals (2.4%); no-screening arm: 304 lung cancers detected in 6,612 individuals, 206 lung cancer-related deaths in 6,612 individuals (3.1%); RR 0.76 (0.61–0.94) ^a	Screening arm: number of lung cancers detected NA, 25 lung cancer-related deaths in 1,317 individuals (1.9%); no-screening arm: 36 lung cancer-related deaths in 1,277 individuals (2.8%); RR 0.67 (0.38–1.14) ^a
LUSI ⁶	7 years after baseline	Screening arm: 43 lung cancers detected among 1,315 individuals (3.3%), 18 lung cancer-related deaths among 1,315 individuals (1.4%); no-screening arm: 19 lung cancer-related deaths among 1,307 individuals (1.5%); HR 0.94 (95% CI 0.54–1.61) ^a	Screening arm: 20 lung cancers detected among 714 individuals (2.8%), 2 lung cancer-related deaths among 714 individuals (0.3%); no screening arm: 10 lung cancer-related deaths among 716 individuals (1.4%); HR 0.31 (95% CI 0.10–0.96) ^a

HR, hazard ratio; NA, not applicable; RR, rate ratio. ^aValues not adjusted for person years.

in women than in men, with RRs of 0.46 (95% CI 0.21–0.96) versus 0.79 (95% CI 0.60–1.03) at 7 years, 0.41 (95% CI 0.19–0.84) versus 0.76 (95% CI 0.60–0.97) at 8 years, and 0.52 (95% CI 0.28–0.94) versus 0.76 (95% CI 0.61–0.96) at 9 years. At 11 years, and thus 5.5 years after the last screening round, the RR was 0.78 overall, indicating the importance of repeated screening and the length of screening intervals.

In the LUSI trial, the difference in lung cancer mortality between individuals in the LDCT screening and no screening arms was not statistically significant (HR 0.74, 95% CI 0.46–1.19; $P=0.21$), possibly owing to the small size of the intervention population⁶ (TABLE 2). However, lung cancer-specific mortality was significantly lower in the screening arm when considering women alone (HR 0.31, 95% CI 0.10–0.96; $P=0.04$).

Taken together, the sex-specific subgroup analyses of the NLST, NELSON and LUSI trials suggest that lung cancer screening could have a more beneficial effect in women than in men, with trends towards fewer late-stage cancers and fewer lung cancer-related deaths in women undergoing LDCT-based screening. The outcomes of these trials are consistent with estimates of the sensitivity of lung cancer detection and mean preclinical durations established through modelling of the natural history of lung cancer using data from the PLCO trial¹⁰⁴ and other clinical studies. In a Swedish cohort study including >23,000 patients with lung adenocarcinomas (LUADs) or squamous cell carcinomas of the lung, women presented with a better performance status, were younger and were more often never smokers at the time of lung cancer diagnosis compared with men ($P\leq 0.04$). Furthermore, women diagnosed with LUAD had a lower comorbidity burden, tumours of a less advanced stage

and a higher proportion of *EGFR*-mutated tumours than men ($P < 0.001$). When comparing survival outcomes based on tumour stage at the time of detection, lung cancer-specific survival was consistently less favourable for men than for women, with a HR of 0.69 (95% CI 0.63–0.76) for stage IA–IIB LUADs and of 0.94 (95% CI 0.88–0.99) for stage IIIB–IV LUADs¹⁰⁵. Similar results from other large-cohort studies, including a study using data from the Surveillance, Epidemiology, and End Results (SEER) database, have shown a beneficial effect of LDCT-based screening on lung cancer-specific survival in women. An analysis of outcomes involving 24,671 men (51.7%) and 23,035 women (48.3%) from this cohort revealed that 5-year lung cancer-specific survival was significantly worse for men than for women (HR 1.24, 95% CI 1.20–1.28; $P < 0.001$), even after adjusting for age, ethnicity, performance status and smoking status¹⁰⁶. Future studies could help to establish whether the use of different lung cancer screening guidelines for men and women could improve screening performance.

Integrated smoking cessation

Many experts in public health have proposed the integration of smoking cessation interventions within LDCT-based screening programmes in the future. For example, EUPS recommends offering advice on smoking cessation to all current smokers¹⁶. NLST and UKLS provide evidence on the effect of in-trial events on smoking cessation. In the NLST analysis, individuals were significantly more likely to quit smoking if abnormal results had been observed in the previous year's screen ($P < 0.0001$)¹⁰⁷. Differences in smoking prevalence among participants in the NLST were detected up to 5 years after the last screen. Around the same time as

the publication of this analysis of the NLST data, results from the Danish Lung Cancer Screening Trial were published. In this trial, 4,104 participants with a smoking history were randomly assigned to undergo annual LDCT-based screening or no screening. At 5 years, no significant differences in annual smoking status were detected between the LDCT group and the control group¹⁰⁸. In fact, the results of this trial were disappointing because the percentage of ex-smokers in both groups combined significantly increased from 24% at baseline to 37% at year 5 of screening ($P < 0.001$)¹⁰⁸. The findings from the UKLS trial support those from the NLST¹⁰⁹ and are opposed to those from the Danish Lung Cancer Screening Trial. In UKLS, independent of the screening result, smoking cessation rates were 8% (36 of 479 individuals) and 14% (75 of 527) in the control and intervention arms, respectively, 2 weeks after baseline scan results or control assignment, and 21% (79 of 377) versus 24% (115 of 488) up to 2 years after recruitment. Participants with a positive screening result were more likely to quit in the longer term compared with those in the control group ($P = 0.007$) and with those receiving a negative result ($P < 0.001$)¹⁰⁹. This observation raises the question as to whether smoking cessation programmes are only effective in participants requiring an intervention for cancer and suggest that such programmes may not yet be successfully integrated into LDCT-based lung cancer screening — addressing this challenge clearly requires further innovative research. The Yorkshire Lung cancer screening trial has an innovative ongoing initiative to integrating smoking cessation and CT screening²⁰.

Kummer et al.¹¹⁰ have identified different patterns of response to patient participation in screening programmes, both from a psychological and behavioural point of view. Their analysis indicated that the simplistic concept linking smoking cessation with involvement in a CT-based screening programme needs to be reconsidered. These programmes require a more in-depth research agenda to ensure that communication of the screening pathway is designed to promote well-being and motivate positive behavioural change, particularly smoking cessation, ultimately maximizing patient benefit. The fact that lung cancer screening of high-risk participants presents a learning opportunity for smoking cessation should be acknowledged, especially among individuals who receive a positive scan result. Nevertheless, further behavioural research is urgently required to evaluate the optimal strategies for integrating smoking cessation interventions within stratified lung cancer screening, which would lead to further reductions in smoking-related morbidity and mortality.

Cost-effectiveness of lung cancer screening

Any innovative health-care technology — either with a curative or preventative intent — requires appraisal of its added value from health regulators. Owing to budget constraints, decision-makers must consider the economic aspects associated with a new technology, analysing the balance between additional costs and health-care benefits through cost-effectiveness analyses. In some countries, innovations such as lung cancer screening might not be introduced if they are not

considered cost-effective. Therefore, these analyses can be crucial in discussions of national lung cancer screening programmes. In this context, a cost-effectiveness model would compare a theoretical population that is screened — with all its additional costs, savings and health benefits — with the same population in the absence of screening. Health benefits are expressed as LYs or quality-adjusted LYs (QALYs) gained. Although screening does not directly create health benefits per se, it enables the early detection of lung cancer and thus improved treatment options, which can result in health benefits. In these models, input parameters on costs and health benefits are often required to be country specific, while screening-related parameters (such as efficacy, sensitivity and specificity) are based on data from large screening trials. Results are expressed as the incremental cost-effectiveness ratio (ICER), reflecting net costs per QALY or LY gained.

Several cost-effectiveness studies on lung cancer screening have been performed using datasets from various specific clinical studies as an input, while accounting for different scenarios^{111–122} (FIG. 4). Using country-specific thresholds for cost-effectiveness, most studies have demonstrated that lung cancer screening can be cost-effective, with ICERs of US\$15,000–100,000 per QALY gained and CAD\$20,000–62,000 per LY gained. For example, Ten Haaf et al.¹¹⁷ considered a scenario in which participants were assumed eligible for screening if they were aged 55–75 years, had smoked >40 pack-years, and were current smokers or had quit <10 years before the first screen. Based on several simulations, lung cancer screening was considered cost-effective against the threshold of CAD\$50,000 per QALY. In the UK, a cost-effectiveness model was developed and utilized in the UKLS trial¹²³. In addition, the UKLS trial investigators reported an estimated cost of ~£8,500 per QALY gained in individuals undergoing screening, although this value was subject to a number of uncertainties²⁸ as it was only based on the UKLS pilot data.

Both annual and biennial screening programmes have been deemed as potentially cost-effective. Goffin et al.¹¹⁶ specifically compared both strategies in a scenario using the NLST eligibility criteria. They concluded that biennial screening used fewer resources and, although associated with lower gains of LYs, resulted in very similar QALY gains over a time frame of 20 years. These researchers estimated that the ICER of annual screening compared with biennial screening was US\$54,000–4.8 million/QALY gained, which would make biennial screening more cost-effective. However, Ten Haaf et al.¹¹⁷ concluded that annual screening was more cost-effective than biennial screening, although less-intensive screening with longer intervals could also represent a cost-effective approach.

The situation in the USA, where US\$100,000 per QALY is considered cost-effective by the federal health-care system (Centers for Medicare & Medicaid Services), is very different from that in Europe (50,000 euro or US\$55,000) and the UK (£20,000–30,000 or US\$26,000–38,000; FIG. 4). Criss et al.¹¹³ developed four models that showed that the NLST, Centers for Medicare & Medicaid Services and USPSTF screening strategies

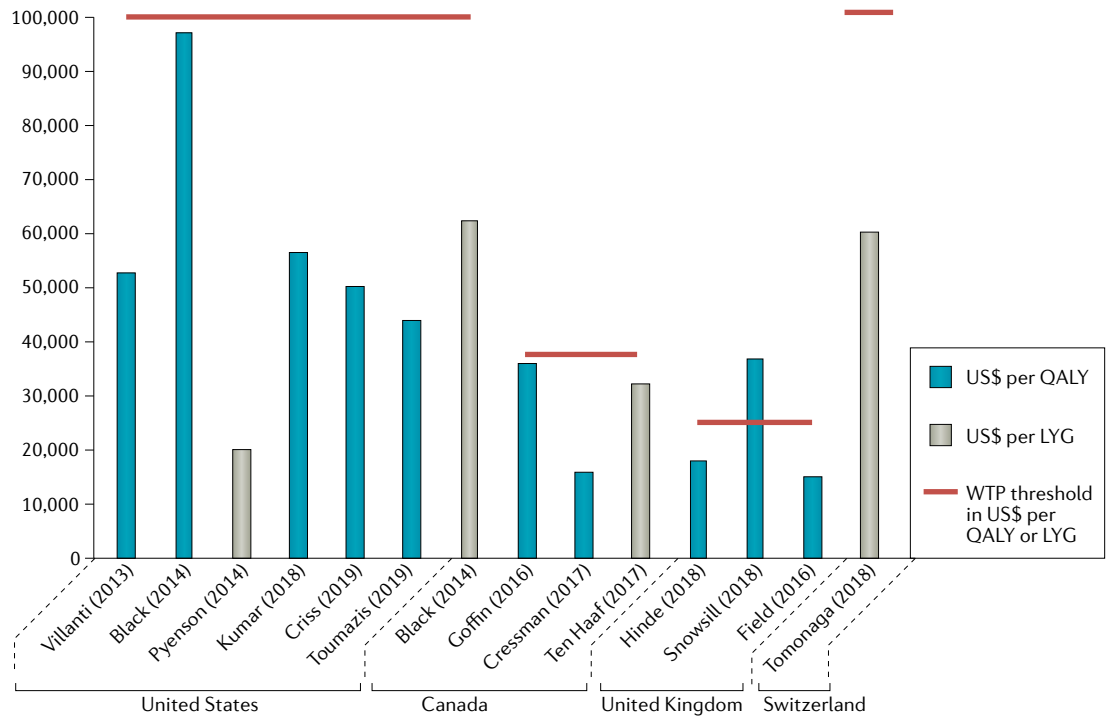


Fig. 4 | Cost-effectiveness of LDCT-based lung cancer screening. Selected published cost-effectiveness analyses of low-dose computed tomography (LDCT)-based lung cancer screening^{111–122}, with results showing incremental cost-effectiveness ratios per quality-adjusted life-years (QALY) or per life-years gained (LYG). The studies are sorted by country and country-specific thresholds for willingness to pay (WTP) are provided. Data are presented in US\$ (with conversion, if necessary) and adjusted to reflect pricing levels in 2019. These studies show that LDCT-based lung cancer screening can be considered cost-effective in most scenarios.

were all cost-effective in the USA, with ICERs averaging US\$49,200, US\$68,600 and US\$96,700 per QALY, respectively. The main difference between these strategies is the maximum age at which to stop screening (80, 77 and 74 years, respectively). This analysis highlighted exactly where the costs lay and the five greatest areas contributing to the total costs associated with screening programmes, noting that the major one is the actual LDCT screening itself. Nevertheless, the major limitation of this analysis was that risk-prediction models for the selection of participants, which could potentially increase the cost-effectiveness of screening, were not factored in. The authors indicated their plan to address this aspect in future projects from the Cancer Intervention and Surveillance Modelling Network. While using a risk-prediction model can increase the cost-effectiveness of a screening programme, related issues that have not been investigated in this context include the tendency of the target population to have comorbidities and, therefore, a shorter life expectancy and potentially a lower quality of life. The latest publication of results from the NELSON trial warrants new cost-effectiveness analyses to assess the financial implications of volumetric-based lung screening¹⁰. The increased availability of data from patients with lung cancer and, in particular, from screening programmes, will make future cost-effectiveness analyses more robust and therefore better suited to assist decision-makers on designing and introducing LDCT-based lung cancer screening in national programmes. Future

cost-effectiveness models could encompass multiple perspectives, such as the health-care and societal perspectives, as well as a fiscal perspective to better determine the financial implications of introducing national lung cancer screening programmes. Future cost-effectiveness models should also consider the costs of expensive targeted agents and immune-checkpoint inhibitors.

Current opportunities worldwide

Screening in China. Lung cancer has been the leading cause of cancer-related death in China since 2005, with an age-standardized 5-year survival of only 19.7% in 2015 (REF.¹²⁴). Data from the National Central Cancer Registry of China in 2014 revealed that, on average, >10,400 lung cancers were diagnosed daily and >6,200 lung cancer-related deaths occurred each day¹²⁵. Lung cancer mortality in China has been projected to increase by ~40% between 2015 and 2030 (REF.¹²⁶). Compared with countries in Europe and North America, in most Asian countries, lung cancer is more frequent even in non-smokers¹²⁷, suggesting that Asian countries might need to use lung cancer screening guidelines different from those we have discussed in the previous sections.

One of the earliest lung cancer screening programmes in China was initiated in 2009 and involved a rural population in the Yunnan Province¹²⁸. Since 2012, the Ministries of Finance and Health of China have included lung cancer screening in the national cancer early detection and treatment programme for the urban population¹²⁷. A modelling study revealed

that LDCT-based screening in urban areas of China would lead to a 17.2% and 24.2% reduction of lung cancer-related mortality compared with chest radiography-based screening and no screening, respectively¹²⁹. In Shanghai, a total of 14,506 individuals were involved in an LDCT-based lung cancer screening study¹³⁰. The pre-set positive result of screening was defined as nodules of any size and any density. The lung cancer detection and incidental detection (that is, detection of any abnormality other than lung cancer) rates were 29.9% and 1.2%, respectively, with an incidental detection rate of stage I lung cancer of 0.97%. The frequency of detection of nodules with a diameter of <5 mm was 74.9%, although 94.1% of lung cancers detected were ≥5 mm, and the frequency of detection of non-solid nodules was 84.9%. Therefore, the baseline LDCT-based lung cancer screening round revealed that subsolid nodules accounted for the majority of lung cancers in the study population and that a diameter of 5 mm is the recommended threshold for positive results¹³⁰.

LDCT-based lung cancer screening has gained popularity in China; however, the definition of the high-risk population and the high number of false-positive results remain two challenges that need to be addressed. Previous studies have shown that the criteria used in Europe and North America to determine individuals at a high risk of lung cancer might not be suitable for the Chinese population, especially considering the high incidence of lung cancer in women and non-smokers in China¹³¹. Optimization of the eligibility criteria and identification of (new) risk factors associated with lung nodule detection are crucial aspects for improving the sensitivity and specificity of LDCT-based lung cancer screening in China. The definition of high-risk criteria in the screening population will depend on the results of future and ongoing multicentre RCTs. Considering the geographic and lifestyle variations across the country, specific high-risk criteria for the major regions might need to be proposed to account for differences in external high-risk factors such as exposure to air pollution in the afternoon, radon (indoors), kitchen fumes and second-hand smoke. Family history and genetic susceptibility should also be considered. Identifying subpopulations at high risk of lung cancer should be a clear priority in China, because no large epidemiological datasets have thus far been used to assess the risk parameters for screening eligibility.

The challenge posed by the high-number of false-positive results is mainly caused by cultural perceptions. In our experience (SY.L.), the medical environment of China tends to favour cautiousness from both clinicians and patients, which could result in overtreatment. A large number of small or intermediate-sized (<5 mm) lung nodules that are detected in >75% of all participants were benign¹³⁰; however, this result increases apprehension in the general population. Currently, the number of nodules detected with a diameter of <3 mm is increasing, especially with the development of AI-based approaches; even for these small nodules, in practice, invasive treatment is often preferred over watchful waiting.

An extensive Review of lung cancer screening in China published in 2019 (REF.¹²⁷) demonstrated a great deal of

lung cancer screening activity throughout the country. However, most of these programmes have reported only preliminary results, mainly through websites and meeting abstracts, and thus the available data need to be interpreted cautiously. The authors of this Review have reported that 23 lung cancer screening programmes have been completed or are ongoing in China since the 1980s, mainly after the year 2000 (REF.¹²⁷). Of note, the entry criteria are generally not smoking stringent owing to the existence of different subpopulations with high risk of lung cancer in China. In this country, the evidence for LDCT-based screening implementation is mainly based on results of RCTs conducted elsewhere. Looking into the future, LDCT-based screening programmes incorporating smoking cessation would result in greater benefits for participants. The recommendations advocated in this extensive Review of lung cancer screening in China are pertinent to future success and need to be implemented¹²⁷ (Supplementary Table 1). Further research in China, where lung cancer is now considered an epidemic, is urgently required.

Screening in Japan and South Korea. To date, few studies have reported on the efficacy of LDCT-based lung cancer screening in non-smokers and light smokers¹³². In Japan, one such study was initiated in the Hitachi district, which included a large proportion (~30%) of individuals aged 50–64 years with a smoking history of <30 pack-years^{133,134}. Lung cancer mortality in this district following screening was found to differ significantly with that in the whole of Japan (2005–2009), with a standardized mortality ratio of 0.76 (95% CI 0.67–0.86; $P < 0.001$). In women, the reduction in standardized mortality ratio was also significant (0.74, 95% CI 0.56–0.97). Of note, ≥90% of women were non-smokers¹³³; these results suggest that LDCT-based screening can lead to a decline in lung cancer-related mortality in both non-smokers and smokers, although the authors identified a number of limitations in this study such as in the trial design, with CT scans only being performed in years 1 and 6.

The National Korean Lung Cancer Screening Project (K-LUCAS) is a single-arm trial aimed at a high-risk population of individuals¹³⁵. The pilot study included 256 individuals, and its purpose was to assess the feasibility of a multicentre nationwide programme using the K-LUCAS protocol¹³⁵. The inclusion criteria for K-LUCAS were the same as those for the NLST. In a pilot test of this trial involving 256 participants, 10 nodules classified as grade 3 according to Lung-RADS were identified, nine grade 4 nodules were identified, and one participant was diagnosed with lung cancer. In addition, 86.3% of participants said they would participate in future lung cancer screening programmes and the average degree of willingness to quit smoking among current smokers was 12.7% higher than before screening.

Future opportunities using AI

The implementation of large lung cancer screening programmes has led to a massive increase in the workload of radiologists¹³⁶. In parallel, technical improvements in LDCT have enabled small-sized pulmonary nodules to be visualized. Over the past decades, efforts have been

made to improve screening procedures using AI-based strategies to detect and classify pulmonary nodules. Before these algorithms are implemented in routine clinical care, their performance should be proven to be robust in external datasets.

Computer-aided detection systems. Different CAD systems have been developed to assist radiologists in identifying relevant nodules. However, the use of CAD remains challenging. A volumetric chest LDCT scan contains >9 million voxels. A lung nodule with a diameter of 5 mm occupies ~130 voxels or only 1.4×10^{-5} of the lung volume¹³⁷. False-negative results (when a clinically significant nodule is not detected) and especially false-positive findings can be common; adding the result of CAD-based assessment to that of a radiologist led to a significantly better performance than that from combining two CAD systems without a human reader (97–99% versus 85–88%; $P < 0.03$)¹³⁸.

The effect of CAD as a second reader has been studied in different LDCT-based lung cancer screening trials. Within a subset of 400 patients from the NELSON trial that had been double-read by radiologists, 22% of nodules $\geq 50 \text{ mm}^3$ were identified solely by CAD, including

one lung cancer¹³⁹. Liang et al. showed that four different CAD systems enabled the identification of 56–70% of 50 tumours (with a mean diameter of 4.8 mm) that had been missed on the prevalence round of the I-ELCAP study but failed to identify 20% of lung cancers identified by radiologists¹⁴⁰. These results suggest that CAD has potential value as a second reader in LDCT-based lung cancer screening, although this approach is currently not part of routine clinical care. The detection rate of current standard LDCT was evaluated using maximum intensity projection (a type of CAD) or two different CAD systems. These systems were associated with comparable incremental sensitivity, with reporting times and false-positive rates favouring maximum intensity projection^{141,142}. However, unlike the capabilities of radiologists, CAD systems keep being substantially improved over time owing to advances in neural network and AI systems^{141,143} and, thus, these systems might have a role in lung cancer screening in the future.

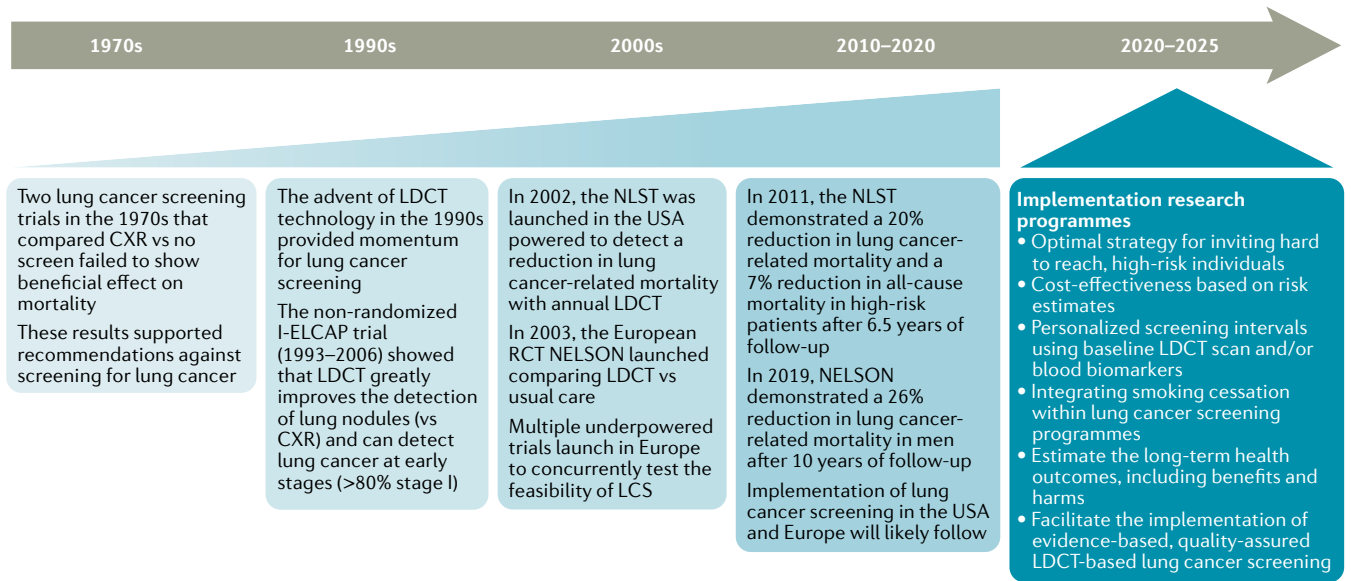
Lung nodule classification. Deep learning (DL)-based approaches can help to accurately distinguish benign from malignant lung nodules as reported in two large-cohort studies published in 2019 (REFS^{144,145}).

Box 1 | Guiding principles for implementation research in lung cancer screening

- Optimize low-dose computed tomography (LDCT)-based lung cancer screening by evaluating the cost-effectiveness of existing as well as innovative approaches based on risk estimates.
- Identify the optimal and most cost-effective strategy for inviting high-risk individuals (often hard to reach) to participate in population-focused LDCT-based lung cancer screening interventions.
- Estimate the effect of personalized, less-intensive screening regimens with longer screening intervals (for example, biennial) relying on health-related risk factors detected on a baseline LDCT scan and, potentially, on blood-based biomarker assays.
- Explore the effect of integrating effective smoking cessation interventions within lung cancer screening programmes.
- Estimate the long-term health outcomes, including benefits and harms, as well as cost-effectiveness by incorporating the above recommendations.

- Develop guidelines and training programmes to facilitate the implementation of evidence-based, quality-assured LDCT-based lung cancer screening as well as appropriate management of LDCT screen-detected pulmonary nodules.

Fifty-year timeline of lung cancer LDCT-based screening and implementation planning. Evidence-based lung cancer screening trials started in the 1970s with trials of chest radiography (CXR) and continued with the pioneering work in LDCT undertaken in the International Early Lung Cancer Action Project (I-ELCAP), the US National Lung Screening Trial (NLST) and seven pilot trials in Europe. Recruitment for the NELSON trial recruitment, the only fully powered lung cancer LDCT-based screening trial in Europe, started in 2003. These five decades of research have now provided the lung cancer community with an international framework for the implementation of lung cancer LDCT-based screening. RCT, randomized controlled trial.



Ardila et al. estimated lung cancer risk with a DL approach mainly based on changes in nodule volume¹⁴⁴. The training set and test set included data from 42,290 and 6,716 NLST participants, respectively, and the algorithm was validated retrospectively in an independent clinical dataset including 1,139 individuals¹⁴⁴. Lung cancer risk estimation was restricted to 1 year after LDCT. For the 6,716 participants (including 86 with cancer) in the test set, the model achieved an AUC of 94.4% (95% CI 91.1–97.3%). A similar result was achieved in the external validation set (1,139 individuals, 27 with cancer), with an AUC of 95.5% (95% CI 83.1–98.0%). Huang et al.¹⁴⁵ focused on nodule classification at the annual follow-up scan rather than at the baseline LDCT scan. Using a DL algorithm (referred to as DeepLR), they identified nodule features predictive of malignancy. For the training set, they used baseline and follow-up LDCT data from 25,097 NLST participants who had undergone at least two LDCT scans. DeepLR was validated in 2,294 participants from the PanCan study; among this high-risk population, the algorithm enabled the identification of a low-risk group (55%) with an estimated probability of developing lung cancer in the following 2 years of only 0.2%. DeepLR outperformed Lung-RADS in predicting lung cancer-related mortality risk (HR 16.07, 95% CI 10.15–25.44; $P < 0.0001$). In addition, DeepLR was associated with a very high true-negative nodule rate, which could enable the potential identification of individuals who would benefit from repeat screening every 2–3 years as opposed to the current recommendation of annual screening¹⁴⁵.

Baldwin et al.¹⁴⁶ compared the performance of an AI-based algorithm, the Lung Cancer Prediction Convolutional Neural Network (LCP-CNN), with that of the Brock parsimonious model in discriminating between benign and malignant pulmonary nodules. Three radiology datasets from the UK were used in this analysis, which revealed AUCs of 89.6% and 86.8% for the LCP-CNN and the Brock parsimonious model, respectively ($P \leq 0.005$). The percentage of nodules with a score

below the lowest category for cancer, and thus not requiring short-term follow-up, were 24.5% and 10.9%, respectively. Of note, this study was performed on a clinical trial dataset with a lung cancer prevalence of 19.3%, which is in contrast with the prevalence typical in lung cancer screening settings (1–3%) and therefore the performance of LCP-CNN in a screening setting is currently unknown.

AI also has the potential to enable the discrimination of different types of lung nodules. A total of 12,754 thin-section chest LDCT scans were retrospectively collected for training, validation and testing of a DL-based convolutional neural network. Pulmonary nodules from these scans were categorized into four types: solid, subsolid, calcified and pleural. The DL model enabled the detection of most nodules when choosing a low-specificity standard. This model had a sensitivity of 99.57 (95% CI 98.62–100.00) and a specificity of 28.03 (95% CI 25.51–30.62) compared with 97.44 (95% CI 95.26–99.18) and 29.23 (95% CI 26.69–31.88), respectively, using the Brock parsimonious model. The success of this model relied on the combination of two convolutional neural network structures¹⁴⁷.

Conclusions

The results from several RCTs of LDCT-based lung cancer screening, including NELSON, have now provided conclusive evidence of a mortality reduction associated with the implementation of lung cancer screening in individuals from both sexes deemed to be at high risk of lung cancer^{10,148}. The lung cancer community now has the opportunity to focus on implementation research, guided by objectives that we have identified thanks to advances over the past decade (BOX 1). The results of these research programmes will help to consolidate international opinion and guide national policy-makers in designing the most appropriate lung cancer screening programmes that are cost-effective for their own diverse health-care systems.

Published online: 12 October 2020

- Bray, F. et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **68**, 394–424 (2018).
- Camidge, D. R., Doebele, R. C. & Kerr, K. M. Comparing and contrasting predictive biomarkers for immunotherapy and targeted therapy of NSCLC. *Nat. Rev. Clin. Oncol.* **16**, 341–355 (2019).
- Goldstraw, P. et al. International Association for the Study of Lung Cancer Staging and Prognostic Factors Committee, Advisory Boards, and Participating Institutions; International Association for the Study of Lung Cancer Staging and Prognostic Factors Committee Advisory Boards and Participating Institutions. The IASLC Lung Cancer Staging Project: Proposals for Revision of the TNM Stage Groupings in the Forthcoming (Eighth) Edition of the TNM Classification for Lung Cancer. *J. Thorac. Oncol.* **11**, 39–51 (2016).
- Aberle, D. R. et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N. Engl. J. Med.* **365**, 395–409 (2011).
- Koning, H. D. E., Van Der Aalst, C. & K. Ten Haaf, M. Oudkerk, PL02.05 effects of volume CT lung cancer screening: mortality results of the NELSON Randomised-Controlled Population Based Trial. *J. Thorac. Oncol.* **13**, S185 (2018).
- Becker, N. et al. Lung cancer mortality reduction by LDCT screening — Results from the randomized German LUSI trial. *Int. J. Cancer* **146**, 1503–1513 (2020).
- Pastorino, U. et al. Prolonged lung cancer screening reduced 10-year mortality in the MILD trial: new confirmation of lung cancer screening efficacy. *Ann. Oncol.* **30**, 1162–1169 (2019).
- Infante, M. et al. Long-term follow-up results of the DANTE trial, a randomized study of lung cancer screening with spiral computed tomography. *Am. J. Respir. Crit. Care Med.* **191**, 1166–1175 (2015).
- Paci, E. et al. Mortality, survival and incidence rates in the ITALUNG randomised lung cancer screening trial. *Thorax* **72**, 825–831 (2017).
- de Koning, H. J. et al. Oudkerk, reduced lung-cancer mortality with volume CT screening in a randomized trial. *N. Engl. J. Med.* <https://doi.org/10.1056/nejmoa1911793> (2020).
- Field, J. K. et al. UK Lung Cancer RCT Pilot Screening Trial: baseline findings from the screening arm provide evidence for the potential implementation of lung cancer screening. *Thorax* **71**, 161–170 (2016).
- Tammemägi, M. C. et al. Participant selection for lung cancer screening by risk modelling (the Pan-Canadian Early Detection of Lung Cancer [PanCan] study): a single-arm, prospective study. *Lancet Oncol.* **18**, 1523–31. (2017).
- Henschke, C. I. et al. Survival of patients with stage I lung cancer detected on CT screening. *N. Engl. J. Med.* **355**, 1763–1771 (2006).
- Villeneuve, P. J. Lifetime probability of developing lung cancer, by smoking status, Canada. *Can. J. Public Health* **85**, 385–388 (1994).
- Bruder, C. et al. Estimating lifetime and 10-year risk of lung cancer. *Prev. Med. Reports* **11**, 125–130 (2018).
- Oudkerk, M. et al. European position statement on lung cancer screening. *Lancet Oncol.* **18**, e754–e766 (2017).
- Ghimire, B. et al. Evaluation of a health service adopting proactive approach to reduce high risk of lung cancer: The Liverpool Healthy Lung Programme. *Lung Cancer* **134**, 66–71 (2019).
- Crosbie, P. A. et al. Implementing lung cancer screening: baseline results from a community-based “Lung Health Check” pilot in deprived areas of Manchester. *Thorax* **74**, 405–409 (2019).
- Bartlett, E. C. et al. Baseline results of the West London lung cancer screening pilot study – Impact of mobile scanners and dual risk model utilisation. *Lung Cancer* **148**, 12–19 (2020).
- ISRCTN Registry The Yorkshire Lung Screening Trial- ISRCTN42704678. <https://doi.org/10.1186/ISRCTN42704678> (2019).
- Raji, O. Y. et al. Predictive accuracy of the Liverpool Lung Project risk model for stratifying patients for computed tomography screening for lung cancer: a case-control and cohort validation study. *Ann. Intern. Med.* **157**, 242–250 (2012).

22. Ten Haaf, K. et al. Risk prediction models for selection of lung cancer screening candidates: A retrospective validation study. *PLoS Med.* **14**, e1002277 (2017).
23. Tammemagi, M. C. et al. Development and validation of a multivariable lung cancer risk prediction model that includes low-dose computed tomography screening results. *JAMA Netw. Open* **2**, e190204. (2019).
24. Cassidy, A. et al. The LLP risk model: an individual risk prediction model for lung cancer. *Br. J. Cancer.* **98**, 270–276 (2008).
25. Tammemagi, M. C. et al. Selection criteria for lung-cancer screening. *N. Engl. J. Med.* **368**, 728–736 (2013).
26. Kovalchik, S. A. et al. Targeting of low-dose CT screening according to the risk of lung-cancer death. *N. Engl. J. Med.* **369**, 245–254 (2013).
27. Katki, H. A., Kovalchik, S. A., Berg, C. D., Cheung, L. C. & Chaturvedi, A. K. Development and validation of risk models to select ever-smokers for ct lung cancer screening. *JAMA* **315**, 2300–2311 (2016).
28. Field, J. K. et al. The UK lung cancer screening trial: a pilot randomised controlled trial of low-dose computed tomography screening for the early detection of lung cancer. *Health Technol. Assess.* **20**, 1–146 (2016).
29. US-National-Library-Medicine. *ClinicalTrials.gov* <https://www.clinicaltrials.gov/ct2/show/NCT03934866> (2018).
30. NHS-England National-Cancer-Programme. Targeted Screening for Lung Cancer with Low Radiation Dose Computed Tomography <https://www.england.nhs.uk/wp-content/uploads/2019/02/targeted-lung-health-checks-standard-protocol-v1.pdf> (2019).
31. Ten Haaf, K. et al. A comparative modeling analysis of risk-based lung cancer screening strategies. *J. Natl Cancer Inst.* **112**, 466–479 (2020).
32. Gordon, W. J., Landman, A., Zhang, H. & Bates, D. W. Beyond validation: getting health apps into clinical practice. *NPJ Digit. Med.* <https://doi.org/10.1038/s41746-019-0212-z> (2020).
33. Seijo, L. M. et al. Biomarkers in lung cancer screening: achievements, promises, and challenges. *J. Thorac. Oncol.* **14**, 343–357 (2019).
34. Amos, C. I., Brennan, P. J., Hung, R. J. & Lin, X. Integrative analysis of lung cancer etiology and risk. *Grantome.com* <https://grantome.com/grant/NIH/U19-CA203654-01A1> (2017).
35. Kachuri, L. et al. Immune-mediated genetic pathways resulting in pulmonary function impairment increase lung cancer susceptibility. *Nat. Commun.* <https://doi.org/10.1038/s41467-019-13855-2> (2020).
36. Liu, Y. et al. Rare variants in known susceptibility loci and their contribution to risk of lung cancer. *J. Thorac. Oncol.* **13**, 1483–1495 (2018).
37. Ji, X. et al. Identification of susceptibility pathways for the role of chromosome 15q25.1 in modifying lung cancer risk. *Nat. Commun.* **9**, 3221 (2018).
38. Guida, F. et al. Assessment of lung cancer risk on the basis of a biomarker panel of circulating proteins. *JAMA Oncol.* **4**, e182078 (2018).
39. Accomasso, L., Cristallini, C. & Giachino, C. Risk assessment and risk minimization in nanomedicine: A need for predictive, alternative, and 3Rs strategies. *Front. Pharmacol.* **9**, 228 (2018).
40. Pham, D., Bhandari, S., Pinkston, C., Oechsli, M. & Kloecker, G. Lung cancer screening registry reveals low-dose CT screening remains heavily underutilized. *Clin. Lung Cancer* **21**, e206–e211 (2020).
41. van der Aalst, C. M., Ten Haaf, K. & de Koning, H. J. Lung cancer screening: latest developments and unanswered questions. *Lancet Respir. Med.* **4**, 749–761 (2016).
42. Ali, N. et al. Barriers to uptake among high-risk individuals declining participation in lung cancer screening: a mixed methods analysis of the UK lung cancer screening (UKLS) trial. *BMJ Open* **5**, e008254 (2015).
43. McRonald, F. E. et al. The UK Lung Screen (UKLS): Demographic profile of first 88,897 approaches provides recommendations for population screening. *Cancer Prev. Res.* **7**, 362–371 (2014).
44. Cassim, S. et al. Patient and carer perceived barriers to early presentation and diagnosis of lung cancer: a systematic review. *BMC Cancer* **19**, 25 (2019).
45. National Lung Screening Trial Research Team. et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N. Engl. J. Med.* **365**, 395–409 (2011).
46. van Klaveren, R. J. et al. Management of lung nodules detected by volume CT scanning. *N. Engl. J. Med.* **361**, 2221–2229 (2009).
47. Field, J. K., Duffy, S. W., Devaraj, A. & Baldwin, D. R. Implementation planning for lung cancer screening: five major challenges. *Lancet Respir. Med.* **4**, 685–687 (2016).
48. Fontana, R. S. et al. Lung cancer screening: the Mayo program. *J. Occup. Med.* **28**, 746–750 (1986).
49. Oken, M. M. et al. Screening by chest radiograph and lung cancer mortality: The Prostate, Lung, Colorectal, and Ovarian (PLCO) randomized trial. *JAMA* **306**, 1865–1873 (2011).
50. American College of Radiology and Radiological Society of North America. Radiation Dose in X-Ray and CT Exams <https://www.radiologyinfo.org/en/pdf/safety-xray.pdf> (2019).
51. Field, J. K. et al. European randomized lung cancer screening trials: Post NLST. *J. Surg. Oncol.* **108**, 280–286 (2013).
52. Han, D. et al. An update on the European lung cancer screening trials and comparison of lung cancer screening recommendations in Europe. *J. Thorac. Imaging* **34**, 65–71 (2019).
53. American College of Radiology. Lung-RADS Version 1.1 <https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/Lung-Rads> (2019).
54. Smith, R. A. et al. Cancer screening in the United States, 2017: a review of current American Cancer Society guidelines and current issues in cancer screening. *CA Cancer J. Clin.* **67**, 100–121 (2017).
55. Walter, J. E. et al. Occurrence and lung cancer probability of new solid nodules at incidence screening with low-dose CT: analysis of data from the randomised, controlled NELSON trial. *Lancet Oncol.* **17**, 907–916 (2016).
56. Wilson, D. O. et al. The Pittsburgh Lung Screening Study (PLuSS): outcomes within 3 years of a first computed tomography scan. *Am. J. Respir. Crit. Care Med.* **178**, 956–961 (2008).
57. Swensen, S. J. et al. CT screening for lung cancer: five-year prospective experience. *Radiology* **235**, 259–265 (2005).
58. van Klaveren, R. J. et al. Management of lung nodules detected by volume CT scanning. *N. Engl. J. Med.* **361**, 2221–2229 (2009).
59. Henschke, C. I. et al. Early lung cancer action project: overall design and findings from baseline screening. *Lancet* **354**, 99–105 (1999).
60. Pedersen, J. H. et al. The Danish randomized lung cancer CT screening trial—overall design and results of the prevalence round. *J. Thorac. Oncol.* **4**, 608–614 (2009).
61. The National Lung Screening Trial Research Team. Results of initial low-dose computed tomographic screening for lung cancer. *N. Engl. J. Med.* **368**, 1980–1991 (2013).
62. Henschke, C. I. et al. Survival of patients with stage I lung cancer detected on CT screening. *N. Engl. J. Med.* **355**, 1763–1771 (2006).
63. Lopes Pegna, A. et al. Design, recruitment and baseline results of the ITALUNG trial for lung cancer screening with low-dose CT. *Lung Cancer* **64**, 34–40 (2009).
64. Becker, N. et al. Randomized study on early detection of lung cancer with MSCT in Germany: study design and results of the first screening round. *J. Cancer Res. Clin. Oncol.* **138**, 1475–1486 (2012).
65. Henschke, C. I. et al. Early lung cancer action project: initial findings on repeat screening. *Cancer* **92**, 153–159 (2001).
66. Swensen, S. J. et al. Screening for lung cancer with low-dose spiral computed tomography. *Am. J. Respir. Crit. Care Med.* **165**, 508–513 (2002).
67. Pinsky, P. F., Gierada, D. S., Hrudaya Nath, P. & Munden, R. Lung cancer risk associated with new solid nodules in the national lung screening trial. *Am. J. Roentgenol.* **209**, 1009–1014 (2017).
68. Saghir, Z. et al. CT screening for lung cancer brings forward early disease. The randomised Danish Lung Cancer Screening Trial: status after five annual screening rounds with low-dose CT. *Thorax* **67**, 296–301 (2012).
69. Heuvelmans, M. A. et al. Relationship between nodule count and lung cancer probability in baseline CT lung cancer screening: The NELSON study. *Lung Cancer* **113**, 45–50 (2017).
70. Walter, J. E. et al. Relationship between the number of new nodules and lung cancer probability in incidence screening rounds of CT lung cancer screening: The NELSON study. *Lung Cancer* **125**, 103–108 (2018).
71. MacMahon, H. et al. Guidelines for management of incidental pulmonary nodules detected on CT images: from the Fleischner Society 2017. *Radiology* **284**, 228–243 (2017).
72. American College of Radiology. LungRADSTM Version 1.0 Assessment Categories Release date: April 28, 2014 https://www.acr.org/-media/ACR/Files/RADS/Lung-RADS/LungRADS_AssessmentCategories.pdf (2014).
73. Han, D. et al. Influence of lung nodule margin on volume- and diameter-based reader variability in CT lung cancer screening. *Br. J. Radiol.* **91**, 20170405 (2018).
74. Heuvelmans, M. A. et al. Disagreement of diameter and volume measurements for pulmonary nodule size estimation in CT lung cancer screening. *Thorax* **73**, 779–781 (2018).
75. Callister, M. E. J. et al. British Thoracic Society guidelines for the investigation and management of pulmonary nodules. *Thorax* **70**, 1–54 (2015).
76. Heuvelmans, M. A. & Oudkerk, M. Pulmonary nodules measurements in CT lung cancer screening. *J. Thorac. Dis.* **10**, S2100–S2102 (2018).
77. Han, D. et al. Influence of lung nodule margin on volume- and diameter-based reader variability in CT lung cancer screening. *Br. J. Radiol.* **91**, 20170405 (2018).
78. Bankier, A. A. et al. Recommendations for measuring pulmonary nodules at CT: a statement from the Fleischner society. *Radiology* **285**, 584–600 (2017).
79. Kakinuma, R. et al. Solitary pure ground-glass nodules 5 mm or smaller: frequency of growth. *Radiology* **276**, 873–882 (2015).
80. Horeweg, N. et al. Lung cancer probability in patients with CT-detected pulmonary nodules: A prespecified analysis of data from the NELSON trial of low-dose CT screening. *Lancet Oncol.* **15**, 1332–1341 (2014).
81. Henschke, C. I. et al. CT screening for lung cancer: part-solid nodules in baseline and annual repeat rounds. *Am. J. Roentgenol.* **207**, 1176–1184 (2016).
82. Yip, R. et al. Lung cancer deaths in the national lung screening trial attributed to nonsolid nodules. *Radiology* **281**, 589–596 (2016).
83. Walter, J. E. et al. New subsolid pulmonary nodules in lung cancer screening: the NELSON trial. *J. Thorac. Oncol.* **13**, 1410–1414 (2018).
84. Henschke, C. I. et al. ELCAP Group. CT screening for lung cancer: frequency and significance of part-solid and nonsolid nodules. *Am. J. Roentgenol.* **178**, 1053–1057 (2002).
85. Lee, S. M. et al. Transient part-solid nodules detected at screening thin-section CT for lung cancer: comparison with persistent part-solid nodules. *Radiology* **255**, 242–251 (2010).
86. Oh, J.-Y. et al. Clinical significance of a solitary ground-glass opacity (GGO) lesion of the lung detected by chest CT. *Lung Cancer* **55**, 67–73 (2007).
87. Kakinuma, R. et al. Natural history of pulmonary subsolid nodules: a prospective multicenter study. *J. Thorac. Oncol.* **11**, 1012–1028 (2016).
88. McWilliams, A. et al. Probability of cancer in pulmonary nodules detected on first screening CT. *N. Engl. J. Med.* **369**, 910–919 (2013).
89. Tammemagi, M. et al. Predicting malignancy risk of screen-detected lung nodules—mean diameter or volume. *J. Thorac. Oncol.* **14**, 203–211 (2019).
90. Marcus, M. W. et al. Probability of cancer in lung nodules using sequential volumetric screening up to 12 months: the UKLS trial. *Thorax* **74**, 761–767 (2019).
91. Raghu, V. K. et al. Feasibility of lung cancer prediction from low-dose CT scan and smoking factors using causal models. *Thorax* **74**, 643–649 (2019).
92. Swensen, S. J., Silverstein, M. D., Ilstrup, D. M., Schleck, C. D. & Edell, E. S. The probability of malignancy in solitary pulmonary nodules. Application to small radiologically indeterminate nodules. *Arch. Intern. Med.* **157**, 849–855 (1997).
93. Li, Y., Chen, K. Z. & Wang, J. Development and validation of a clinical prediction model to estimate the probability of malignancy in solitary pulmonary nodules in Chinese people. *Clin. Lung Cancer* **12**, 313–319 (2011).
94. González Maldonado, S. et al. Evaluation of prediction models for identifying malignancy in pulmonary nodules detected via low-dose computed tomography. *JAMA Netw. Open* **3**, e1921221 (2020).
95. Heuvelmans, M. A. & Oudkerk, M. Less is more in lung cancer risk prediction models. *JAMA Netw. Open* **3**, e1921492 (2020).
96. De Koning, H. J. et al. Benefits and harms of computed tomography lung cancer screening strategies: a comparative modeling study for the U.S. Preventive services task force. *Ann. Intern. Med.* **160**, 311–320 (2014).

97. Patz, E. F. et al. Lung cancer incidence and mortality in National Lung Screening Trial participants who underwent low-dose CT prevalence screening: a retrospective cohort analysis of a randomised, multicentre, diagnostic screening trial. *Lancet Oncol.* **17**, 590–599 (2016).
98. Pastorino, U. et al. Ten-year results of the Multicentric Italian Lung Detection trial demonstrate the safety and efficacy of biennial lung cancer screening. *Eur. J. Cancer* **118**, 142–148 (2019).
99. Schreuder, A. et al. Lung cancer risk to personalise annual and biennial follow-up computed tomography screening. *Thorax* <https://doi.org/10.1136/thoraxjnl-2017-211107> (2018).
100. Yousaf-Khan, U. et al. Final screening round of the NELSON lung cancer screening trial: the effect of a 2.5-year screening interval. *Thorax* **72**, 1–9 (2016).
101. Heuvelmans, M. A. & Oudkerk, M. Appropriate screening intervals in low-dose CT lung cancer screening. *Transl Lung Cancer Res.* **7**, 281–287 (2018).
102. Pinsky, P. F., Church, T. R., Izmirlian, G. & Kramer, B. S. The National Lung Screening Trial: results stratified by demographics, smoking history, and lung cancer histology. *Cancer* **119**, 3976–3983 (2013).
103. National Lung Screening Trial Research Team. Lung cancer incidence and mortality with extended follow-up in the national lung screening trial. *J. Thorac. Oncol.* **14**, 1732–1742 (2019).
104. Ten Haaf, K., Van Rosmalen, J. & de Koning, H. J. Lung cancer detectability by test, histology, stage, and gender: Estimates from the NLST and the PLCO trials. *Cancer Epidemiol. Biomarkers Prev.* **24**, 154–161 (2015).
105. Radkiewicz, C. et al. Sex and survival in non-small cell lung cancer: a nationwide cohort study. *PLoS ONE* **14**, e0219206 (2019).
106. Wainer, Z. et al. Sex-dependent staging in non-small-cell lung cancer; analysis of the effect of sex differences in the eighth edition of the Tumor, Node, Metastases Staging System. *Clin. Lung Cancer* **19**, e933–e944 (2018).
107. Tammemägi, M. C., Berg, C. D., Riley, T. L., Cunningham, C. R. & Taylor, K. L. Impact of lung cancer screening results on smoking cessation. *J. Natl Cancer Inst.* **106**, dju084 (2014).
108. Ashraf, H. et al. Smoking habits in the randomised Danish Lung Cancer Screening Trial with low-dose CT: final results after a 5-year screening programme. *Thorax* **69**, 574–579 (2014).
109. Brain, K. et al. Impact of low-dose CT screening on smoking cessation among high-risk participants in the UK Lung Cancer Screening trial. *Thorax* **72**, 912–918 (2017).
110. Kummer, S. et al. Mapping the spectrum of psychological and behavioural responses to low-dose CT lung cancer screening offered within a Lung Health Check. *Health Expect.* **23**, 433–441 (2020).
111. Villanti, A. C., Jiang, Y., Abrams, D. B. & Pyenson, B. S. A cost-utility analysis of lung cancer screening and the additional benefits of incorporating smoking cessation interventions. *PLoS ONE* **8**, e71379 (2013).
112. Black, W. C. et al. Cost-effectiveness of CT screening in the National Lung Screening Trial. *N. Engl. J. Med.* **371**, 1793–1802 (2014).
113. Criss, S. D. et al. Cost-effectiveness analysis of lung cancer screening in the United States: a Comparative Modeling Study. *Ann. Intern. Med.* **171**, 796–804 (2019).
114. Toumazis, I. et al. Cost-effectiveness analysis of lung cancer screening accounting for the effect of indeterminate findings. *JNCI Cancer Spectr.* **3**, pkz035 (2019).
115. Pyenson, B. S., Henschke, C. I., Yankelevitz, D. F., Yip, R. & Dec, E. Offering lung cancer screening to high-risk medicare beneficiaries saves lives and is cost-effective: an actuarial analysis. *Am. Health Drug Benefits* **7**, 272–282 (2014).
116. Goffin, J. R. et al. Biennial lung cancer screening in Canada with smoking cessation — outcomes and cost-effectiveness. *Lung Cancer* **101**, 98–103 (2016).
117. Ten Haaf, K. et al. Performance and cost-effectiveness of computed tomography lung cancer screening scenarios in a population-based setting: a microsimulation modeling analysis in Ontario, Canada. *PLOS Med.* **14**, e1002225 (2017).
118. Cressman, S. et al. The cost-effectiveness of high-risk lung cancer screening and drivers of program efficiency. *J. Thorac. Oncol.* **12**, 1210–1222 (2017).
119. Tomonaga, Y. et al. Cost-effectiveness of low-dose CT screening for lung cancer in a European country with high prevalence of smoking — a modelling study. *Lung Cancer* **121**, 61–69 (2018).
120. Kumar, V. et al. Risk-targeted lung cancer screening a cost-effectiveness analysis. *Ann. Intern. Med.* **168**, 161–169 (2018).
121. Hinde, S. et al. The cost-effectiveness of the Manchester 'lung health checks', a community-based lung cancer low-dose CT screening pilot. *Lung Cancer* **126**, 119–124 (2018).
122. Snowsill, T. et al. Low-dose computed tomography for lung cancer screening in high-risk populations: a systematic review and economic evaluation. *Health Technol. Assess.* **22**, 1–276 (2018).
123. Whynes, D. K. Could CT screening for lung cancer ever be cost effective in the United Kingdom? *Cost Eff. Resour. Alloc.* **6**, 5 (2008).
124. Zeng, H. et al. Changing cancer survival in China during 2003–15: a pooled analysis of 17 population-based cancer registries. *Lancet Glob. Health* **6**, e555–e567 (2018).
125. Chen, W. et al. Cancer incidence and mortality in China, 2014. *Chin. J. Cancer Res.* **30**, 1–12 (2018).
126. Martín-Sánchez, J. C. et al. Projections in breast and lung cancer mortality among women: a Bayesian analysis of 52 countries worldwide. *Cancer Res.* **78**, 4436–4442 (2018).
127. Cheng, Y. I., Davies, M. P. A., Liu, D., Li, W. & Field, J. K. Implementation planning for lung cancer screening in China. *Precis. Clin. Med.* **2**, 13–44 (2019).
128. Zhou, Q. et al. Demonstration program of population-based lung cancer screening in China: Rationale and study design. *Thorac. Cancer* **5**, 197–203 (2014).
129. Wang, Z. et al. Mortality outcomes of low-dose computed tomography screening for lung cancer in urban China: a decision analysis and implications for practice. *Chin. J. Cancer* **36**, 57 (2017).
130. Fan, L. et al. Lung cancer screening with low-dose CT: baseline screening results in Shanghai. *Acad. Radiol.* **26**, 1283–1291 (2019).
131. Luo, X. et al. Should nonsmokers be excluded from early lung cancer screening with low-dose spiral computed tomography? community-based practice in Shanghai. *Transl Oncol.* **10**, 485–490 (2017).
132. Nawa, T. et al. A decrease in lung cancer mortality following the introduction of low-dose chest CT screening in Hitachi, Japan. *Lung Cancer* **78**, 225–228 (2012).
133. Sagawa, M. et al. Efficacy of low-dose computed tomography screening for lung cancer: the current state of evidence of mortality reduction. *Surg. Today* **47**, 783–788 (2017).
134. Nawa, T. et al. A population-based cohort study to evaluate the effectiveness of lung cancer screening using low-dose CT in Hitachi city, Japan. *Jpn. J. Clin. Oncol.* **49**, 130–136 (2019).
135. Lee, J. et al. Development of protocol for Korean Lung Cancer Screening Project (K-LUCAS) to evaluate effectiveness and feasibility to implement national cancer screening program. *Cancer Res. Treat.* **51**, 1285–1294 (2019).
136. The Royal College of Radiologists. www.rcr.ac.uk/publication/clinical-radiology-uk-workforce-census-report-2018 (2019).
137. Ather, S., Kadir, T. & Gleeson, F. Artificial intelligence and radiomics in pulmonary nodule management: current status and future applications. *Clin. Radiol.* **75**, 13–19 (2020).
138. Christe, A. et al. Lung cancer screening with CT: Evaluation of radiologists and different computer assisted detection software (CAD) as first and second readers for lung nodule detection at different dose levels. *Eur. J. Radiol.* **82**, e873–8 (2013).
139. Wang, Y. et al. No benefit for consensus double reading at baseline screening for lung cancer with the use of semiautomated volumetry software. *Radiology* **262**, 320–326 (2012).
140. Liang, M. et al. Low-dose CT screening for lung cancer: computer-aided detection of missed lung cancers. *Radiology* **281**, 279–288 (2016).
141. Vlahos, I. et al. Lung cancer screening: Nodule identification and characterization. *Transl Lung Cancer Res.* **7**, 288–303 (2018).
142. Ebner, L. et al. Maximum-intensity-projection and computer-aided-detection algorithms as stand-alone reader devices in lung cancer screening using different dose levels and reconstruction kernels. *Am. J. Roentgenol.* **207**, 282–288 (2016).
143. Ciompi, F. et al. Towards automatic pulmonary nodule management in lung cancer screening with deep learning. *Sci. Rep.* **7**, 46479 (2017).
144. Ardila, D. et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat. Med.* **25**, 954–961 (2019).
145. Huang, P. et al. Prediction of lung cancer risk at follow-up screening with low-dose CT: a training and validation study of a deep learning method. *Lancet Digit. Health* **1**, e353–e362 (2019).
146. Baldwin, D. R. et al. External validation of a convolutional neural network artificial intelligence tool to predict malignancy in pulmonary nodules. *Thorax* **75**, 306–312 (2020).
147. Liu, K. et al. Evaluating a fully automated pulmonary nodule detection approach and its impact on radiologist performance. *Radiol. Artif. Intell.* <https://doi.org/10.1148/ryai.2019180084> (2019).
148. Duffy, S. W. & Field, J. K. Mortality reduction with low-dose CT screening for lung cancer. *N. Engl. J. Med.* **382**, 572–573 (2020).

Author contributions

M.O. contributed to all aspects of the preparation of this article. S.L. and J.E.W. researched data. M.A.H. contributed to discussions, researched data and supported the draft of the manuscript. J.K.F. contributed to all aspects of the preparation of this article. The final manuscript was approved by all authors.

Competing interests

The authors declare no competing interests.

Peer review information

Nature Reviews Clinical Oncology thanks S. Lam, U. Pastorino and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41571-020-00432-6>.

© Springer Nature Limited 2020