

# Low-dose computed tomography for lung cancer screening in high-risk populations: a systematic review and economic evaluation

Tristan Snowsill,<sup>1\*</sup> Huiqin Yang,<sup>1</sup> Ed Griffin,<sup>1</sup>  
Linda Long,<sup>1</sup> Jo Varley-Campbell,<sup>1</sup> Helen Coelho,<sup>1</sup>  
Sophie Robinson<sup>1</sup> and Chris Hyde<sup>1,2</sup>

<sup>1</sup>Peninsula Technology Assessment Group (PenTAG), University of Exeter Medical School, Exeter, UK

<sup>2</sup>Exeter Test Group, University of Exeter Medical School, Exeter, UK

\*Corresponding author [t.m.snowsill@exeter.ac.uk](mailto:t.m.snowsill@exeter.ac.uk)

**Declared competing interests of authors:** none

Published November 2018

DOI: 10.3310/hta22690

## Scientific summary

### Low-dose computed tomography for lung cancer screening

Health Technology Assessment 2018; Vol. 22: No. 69

DOI: 10.3310/hta22690

NIHR Journals Library [www.journalslibrary.nihr.ac.uk](http://www.journalslibrary.nihr.ac.uk)

# Scientific summary

## Background

Approximately 46,400 cases of lung cancer were diagnosed in the UK in 2014, representing 13% of the total number of cancer cases. Diagnosis of lung cancer frequently occurs in its later stages. Low-dose computed tomography (LDCT) could detect lung cancer in its early stages, but its clinical effectiveness and cost-effectiveness in a UK national screening setting are uncertain.

## Objectives

This assessment aims to evaluate the clinical effectiveness and cost-effectiveness of using LDCT in screening programmes for lung cancers in high-risk populations.

## Methods

### *Clinical effectiveness*

For the systematic review, a range of bibliographic databases including MEDLINE (via Ovid), MEDLINE In-Process & Other Non-Indexed Citations (via Ovid), EMBASE (via Ovid), PsycINFO (via Ovid), Web of Science (via Clarivate Analytics), CDSR (Cochrane Database of Systematic Reviews), CENTRAL (Cochrane Central Register of Controlled Trials) and CINAHL (Cumulative Index to Nursing and Allied Health Literature) were searched from January 2004 to January 2017. Both published and unpublished literature were identified from systematic searches of electronic sources, consultation with experts in the field and reference checking of relevant systematic reviews.

We included randomised controlled trials (RCTs) involving populations at high risk of lung cancer. Any definitions of high-risk populations were eligible. LDCT screening programmes included both single and multiple rounds. The eligible comparators were no screening or other imaging technology screening programmes [such as chest X-ray (CXR)]. The key outcomes included lung cancer mortality, all-cause mortality, numbers of lung cancers and their stages, health-related quality of life (HRQoL) and smoking behaviour.

Two researchers independently screened the titles and abstracts of all reports identified by the search strategy and full-text papers were subsequently obtained for assessment. Data extraction and quality assessment were undertaken by one researcher and checked by a second. The risk of bias of included studies was assessed using the Cochrane Risk of Bias tool.

Data were tabulated and primarily considered in a narrative review. When appropriate, DerSimonian and Laird random-effects meta-analyses were performed to pool the estimates of effect. A random-effects approach was prespecified as part of the protocol development process; a fixed approach was not favoured as it was thought highly unlikely that only random variation would account for differences between the results of included studies. Statistical heterogeneity was assessed using the  $I^2$ -statistic. We considered the following factors for the exploration of heterogeneity: quality of trials, nature of interventions (e.g. frequency of LDCT screening) and nature of control groups (e.g. best available care such as CXR screening or usual care).

Network meta-analysis was performed to assess the relative effectiveness of three screening strategies (LDCT, usual care and CXR). We estimated the relative ranking probability of each intervention and

obtained the treatment hierarchy of competing interventions using rankogram, surface under the cumulative ranking curve and mean ranks.

### **Cost-effectiveness**

A systematic review of economic evaluations of LDCT screening for lung cancer was undertaken, including a narrative synthesis.

A decision-analytic model was built in Microsoft Excel® 2013 (Microsoft Corporation, Redmond, WA, USA) using a decision tree and discrete event simulation approach. A lifetime time horizon was used, costs (2016 prices) were included from a NHS and Personal Social Services perspective and discounted at 3.5% per year. Health effects on targeted individuals were included and discounted at 3.5% per year.

A natural history model for lung cancer was developed including seven substages of lung cancer development, with progressively worsening survival. The rates of incidence, preclinical progression and clinical presentation were estimated by calibrating the natural history model to data from a large RCT of LDCT screening and to English cancer incidence statistics. LDCT screening was assumed to have imperfect diagnostic performance (i.e. some lung cancers would not be detected) and some individuals without lung cancer would receive further tests.

A risk prediction component was incorporated, so that each individual's short-term risk of lung cancer would be estimated and would be used to decide whether or not the individual was eligible for LDCT screening.

Twelve combinations of age limits and risk thresholds were considered, as were four frequencies of screening: (1) single screen (S), (2) triple screen (T), (3) annual screening (A) and (4) biennial screening (B). Forty-eight screening programmes were evaluated in total and were compared with current practice (no screening).

Costs were estimated from English cost data sets or from the published literature. Health state utility values were estimated from the published literature.

The primary outcome for each strategy was the fully incremental cost-effectiveness ratio (ICER), which is defined for strategies on the cost-effectiveness frontier (strategies that are optimal for some cost-effectiveness threshold) as the ratio of incremental costs to incremental benefits relative to the next worse strategy. Benefits were measured in quality-adjusted life-years (QALYs).

### **Patient and public involvement**

We collected views about the possible impact of a lung cancer screening programme in the UK. We conducted three workshop meetings in which we elicited views from a range of patient and public members, with a particular focus on smokers/former smokers currently without symptoms of lung cancer who would be the target group for any UK national lung cancer screening programme.

## **Results**

### **Clinical effectiveness**

Twelve RCTs were included in the systematic review of clinical effectiveness. Six of these contributed to the key outcomes. Most studies were conducted in European countries but some studies were conducted in the USA, including by far the largest, National Lung Screening Trial (NLST), with > 50,000 participants. One trial, the UK Lung Cancer Screening Trial (UKLS), was conducted in the UK. Most RCTs started between 2001 and 2010, so many are just maturing. The majority of included trials were judged to be of moderate to high quality, but two trials were judged to be of poor quality, including one that contributed mortality data. There was variation between the LDCT programmes, but typically they involved 3–5 rounds of

screening over 3–6.5 years. UKLS, a pilot trial, had only one screening round. The nature of high-risk participants also varied but was usually defined in terms of age and current and past smoking. Of the trials, NLST stands apart, not just in terms of size, but by comparing LDCT to CXR screening rather than no screening.

Concerning mortality, only four of the RCTs, including NLST, currently contribute. Meta-analysis of these showed that LDCT screening was associated with a non-statistically significant decrease in lung cancer mortality [pooled relative risk (RR) 0.94, 95% confidence interval (CI) 0.74 to 1.19] with up to 9.80 years of follow-up when compared with controls (usual care/best available care). A moderate level of heterogeneity was observed ( $I^2 = 43.3\%$ ); therefore, the results should be treated with caution.

A range of potential sources for heterogeneity was investigated. When removing the poor-quality trial [Multicentric Italian Lung Detection (MILD)], sensitivity analysis demonstrated a statistically significant decrease in lung cancer mortality (pooled RR 0.85, 95% CI 0.74 to 0.98) in favour of LDCT screening with considerable reduction in heterogeneity ( $I^2 = 6.9\%$ ).

For all-cause mortality, the review showed a non-statistically significant increase (pooled RR 1.01, 95% CI 0.87 to 1.16) for LDCT screening. Again, given the substantial heterogeneity ( $I^2 = 57.0\%$ ), this pooled estimate should be treated with caution. In the investigation of heterogeneity, removing the low-quality trial (MILD) showed a non-statistically significant decrease in all-cause mortality (pooled RR 0.95, 95% CI 0.89 to 1.00) with considerable reduction in heterogeneity ( $I^2 = 0\%$ ).

Network meta-analysis (including six RCTs) was performed to assess the relative clinical effectiveness of LDCT, usual care and CXR screening. The results showed that LDCT was ranked as the best screening strategy, with a 74.8% probability of being the best intervention in terms of lung cancer mortality reduction. Usual care (no screening) had a 74.7% probability of being the second best strategy and CXR screening a 99.7% probability of being the worst strategy. Both consistency and inconsistency models were fit for lung cancer mortality data. By applying the design-by-treatment model, we did not find any evidence of inconsistency. The global test for inconsistency gives a  $p$ -value of 0.29, indicating no evidence of inconsistency.

Concerning numbers of lung cancers detected, compared with controls (usual care/best available care), LDCT screening was associated with a statistically significant increase (pooled RR 1.38, 95% CI 1.02 to 1.86) with at least 5 years' follow-up. Although there was heterogeneity ( $I^2 = 79.7\%$ ), all included studies individually showed statistically significant increases in the number of cancers detected in the LDCT group. Our findings further demonstrated a shift due to LDCT screening on the stage distribution towards earlier stages for detection of lung cancers. LDCT screening was associated with a statistically significant increase in early-stage (I and II) cancer detection (pooled RR 1.73, 95% CI 1.27 to 2.37) with a corresponding statistically significant decrease in late-stage (III and IV) cancer. There was a statistically significant reduction in the absolute risk of late-stage lung cancer, indicating that there is an element of actual stage shift (pooled RR 0.85, 95% CI 0.73 to 1.00).

Based on the randomised data from four included trials, there were consistently no statistically significant differences in HRQoL or psychological consequences between the LDCT screening groups and control groups during the trials.

The data from three included trials (one reported as two subcomponents) showed mixed results with regard to the effect of a LDCT screening programme on participants' smoking behaviours. The data within trial arms sometimes indicated positive associations between smoking cessation and the presence of an abnormality on LDCT. However, this is inconsistent with the evidence comparing trial arms that did not show a consistent pattern favouring LDCT's effect on smoking behaviour.

### Cost-effectiveness

Existing economic evaluations of LDCT screening for lung cancer have produced markedly variable estimates of the cost-effectiveness of screening. Nineteen studies were identified, with variable quality, and with ICERs ranging from low thousands of US dollars per QALY to in excess of US\$100,000 per QALY.

In the independent economic evaluation (base-case analysis), it was estimated that it would not be cost-effective to screen for lung cancer by LDCT at a cost-effectiveness threshold of £20,000 per QALY, whereas it was estimated that a single screen for individuals aged 60–75 years with at least a 3% risk of lung cancer (S–60–75–3%) would be cost-effective at a cost-effectiveness threshold of £30,000 per QALY (ICER approximately £28,000 per QALY). Three other screening programmes were on the cost-effectiveness frontier; these all used a 3% risk threshold, two were single screens and one was a triple screen. The ICERs for these programmes were > £30,000 per QALY and, therefore, would not be considered cost-effective at commonly used thresholds in the UK. Additionally, annual and biennial screening programmes were not predicted to be cost-effective at any cost-effectiveness threshold.

Screening was predicted to improve the stage distribution and survival of lung cancer, but also to result in overdiagnosis (diagnosis of lung cancers that never would have clinically presented).

A single screen was predicted to reduce lung cancer mortality by 4.2%, and triple screening was predicted to reduce mortality by 4.4%.

A probabilistic sensitivity analysis (PSA) produced similar results to the base-case analysis, although the ICER for S–60–75–3% was higher (approximately £36,000 per QALY). Results from the PSA are typically preferred as they capture non-linear associations between inputs and outcomes.

One-way sensitivity analyses showed that the results were sensitive to the natural history of lung cancer, the cost of lung cancer and the cost of LDCT screening.

Scenario analyses showed that cost-effectiveness was worsened when the natural history model included heterogeneity in the rate of lung cancer progression, and when the mortality effect from screening was attenuated or eliminated. Cost-effectiveness was improved if there was no negative impact on HRQoL from false-positive or indeterminate results.

### Patient and public involvement

An explanatory model was constructed detailing the key associations and core dynamics arising from our patient and public involvement meetings. The model details views around decisions to attend a national lung cancer screening programme, together with views on the broader cultural and societal influences that may affect such decisions. The potential impact of wider societal and cultural contexts on decisions to attend lung cancer screening were discussed. Poor public awareness about potentially effective treatments for lung cancer and survival benefits resulting from early detection was acknowledged, as was a culture of stigma and blame associated with smoking.

### Conclusions

Low-dose computed tomography screening may be clinically effective in reducing lung cancer mortality but there is considerable uncertainty. This arises from the imprecision of pooled estimates, the heterogeneity between the results of the included studies, the fact that the key RCT compares LDCT with CXR screening and the finding from our network meta-analysis that screening with CXR may be associated with worse outcomes than no screening.

Beyond mortality, the review confirms the theoretical basis of LDCT by showing that more lung cancers are diagnosed in the earlier stages and fewer in the later stages. However, it also confirms that more lung

cancers are detected in the LDCT trial arms many years after completion of the screening programmes, indicating an element of overdiagnosis.

It seems unlikely that LDCT screening leads to major differences in psychological consequences and HRQoL, and the effect on smoking behaviour continues to be uncertain.

Evidence from economic modelling suggests that LDCT screening for lung cancer may not be cost-effective, depending on the cost-effectiveness threshold used. Thresholds of £20,000 to £30,000 per QALY are commonly used in the UK, and screening is estimated to be cost-effective with the higher threshold (when using mean values for inputs), but not with the lower. When the probable range of inputs is considered in a PSA, it is estimated that screening is not cost-effective with either threshold.

It is estimated that a national screening programme could result in up to half a million additional computed tomography (CT) screens a year, compared with 2 million CT screens currently conducted each year (in England). It is unlikely such an increase in the burden on radiography services would be accommodated without significant recruitment and/or service reconfiguration.

## Recommendations for research

Clinical effectiveness and cost-effectiveness estimates should be updated with the anticipated results from several ongoing RCTs [particularly the NEderlands Leuven Longkanker Screenings ONderzoek (NELSON) screening trial]. This is likely to resolve many current uncertainties within a reasonable time.

In the longer term, another large trial of lung cancer screening is currently being conducted in Asia that will further explore the generalisability of the initial trial results to populations with different ethnicities.

Further investigation of the quality of currently included trials needs to be conducted to confirm whether they were all truly RCTs or not. If not, they would be appropriately excluded in future systematic reviews on the effectiveness of LDCT.

Further research on why the results of economic evaluations of lung cancer screening vary might enable model selection and model averaging to obtain best estimates of cost-effectiveness while also reflecting structural uncertainty. In addition, certain key costs for the generalisation of economic evaluation results (e.g. the cost of lung cancer) should be estimated from high-quality, representative and recently collected data.

## Registration

This study is registered as PROSPERO CRD42016048530.

## Funding

The National Institute for Health Research Health Technology Assessment programme.

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 4.513

*Health Technology Assessment* is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the Clarivate Analytics Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) ([www.publicationethics.org/](http://www.publicationethics.org/)).

Editorial contact: [journals.library@nhr.ac.uk](mailto:journals.library@nhr.ac.uk)

The full HTA archive is freely available to view online at [www.journalslibrary.nhr.ac.uk/hta](http://www.journalslibrary.nhr.ac.uk/hta). Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: [www.journalslibrary.nhr.ac.uk](http://www.journalslibrary.nhr.ac.uk)

## Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

## HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: <http://www.nets.nhr.ac.uk/programmes/hta>

## This report

The research reported in this issue of the journal was funded by the HTA programme as project number 14/151/07. The contractual start date was in June 2016. The draft report began editorial review in August 2017 and was accepted for publication in February 2018. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care.

**© Queen's Printer and Controller of HMSO 2018. This work was produced by Snowsill *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.**

Published by the NIHR Journals Library ([www.journalslibrary.nhr.ac.uk](http://www.journalslibrary.nhr.ac.uk)), produced by Prepress Projects Ltd, Perth, Scotland ([www.prepress-projects.co.uk](http://www.prepress-projects.co.uk)).

## NIHR Journals Library Editor-in-Chief

**Professor Ken Stein** Chair of HTA and EME Editorial Board and Professor of Public Health, University of Exeter Medical School, UK

## NIHR Journals Library Editors

**Professor Ken Stein** Chair of HTA and EME Editorial Board and Professor of Public Health, University of Exeter Medical School, UK

**Professor Andrée Le May** Chair of NIHR Journals Library Editorial Group (HS&DR, PGfAR, PHR journals)

**Professor Matthias Beck** Professor of Management, Cork University Business School, Department of Management and Marketing, University College Cork, Ireland

**Dr Tessa Crilly** Director, Crystal Blue Consulting Ltd, UK

**Dr Eugenia Cronin** Senior Scientific Advisor, Wessex Institute, UK

**Dr Peter Davidson** Consultant Advisor, Wessex Institute, University of Southampton, UK

**Ms Tara Lamont** Scientific Advisor, NETSCC, UK

**Dr Catriona McDaid** Senior Research Fellow, York Trials Unit, Department of Health Sciences, University of York, UK

**Professor William McGuire** Professor of Child Health, Hull York Medical School, University of York, UK

**Professor Geoffrey Meads** Professor of Wellbeing Research, University of Winchester, UK

**Professor John Norrie** Chair in Medical Statistics, University of Edinburgh, UK

**Professor John Powell** Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), UK

**Professor James Raftery** Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

**Dr Rob Riemsma** Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

**Professor Helen Roberts** Professor of Child Health Research, UCL Great Ormond Street Institute of Child Health, UK

**Professor Jonathan Ross** Professor of Sexual Health and HIV, University Hospital Birmingham, UK

**Professor Helen Snooks** Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

**Professor Jim Thornton** Professor of Obstetrics and Gynaecology, Faculty of Medicine and Health Sciences, University of Nottingham, UK

**Professor Martin Underwood** Warwick Clinical Trials Unit, Warwick Medical School, University of Warwick, UK

Please visit the website for a list of editors: [www.journalslibrary.nihr.ac.uk/about/editors](http://www.journalslibrary.nihr.ac.uk/about/editors)

**Editorial contact:** [journals.library@nihr.ac.uk](mailto:journals.library@nihr.ac.uk)