EVIDENCE FOR DECISION MAKING IN HEALTH CARE

BASEL INSTITUTE FOR CLINICAL EPIDEMIOLOGY AND BIOSTATISTICS

Low-dose CT screening for lung cancer

Final report

Publication date

June 15, 2022

Version

1.2

Authors (alphabetic order)

Dr. Soheila Aghlmandi PhD; Dr. Arjun Bhadhuri, PhD; Prof. Heiner C. Bucher, MD, MPH; Koen de Nijs, MSc; Dr. Hannah Ewald PhD; Dr. sc. Dominik Glinz, PhD, MSc; Dr. Viktoria Gloy, PhD; Alexandra Griessbach, MSc; Dr. David Shaw, PhD, MA, MSc, MML; Ala Taji Heravi, MSc; Dr. Kevin ten Haaf, PhD, MSc; Dr. Yuki Tomonaga, PhD, MSc.

Head of institute

Prof. Dr. med. Heiner C. Bucher, MPH

Universitätsspital Basel

Klinische Epidemiologie

2. Acknowledgement

We thank Diana Grauwiler for administrative support for this report.

3. Contributions

Scoping: DG, HCB; **Systematic review:** HCB and YT wrote the protocol; HE with assistance from HCB developed and conducted the literature search; VG and HCB screened the literature; AG, AB, ATH, YT, and HCB were responsible for data extraction; AG, AB, ATH, YT and HCB graded the risk of bias and the quality of evidence; SA and HCB conducted data analysis; SA, YT and HCB are responsible for data interpretation; HCB and wrote the systematic review section. **Cost-effectiveness and budget impact analysis:** YT wrote the protocol; YT and AB conducted the literature search, screened the literature, performed data extraction and quality assessment; YT summarized the existing literature; KtH and KdN built and calibrated the economic model; YT, KtH, and KdN are responsible for data interpretation of the cost-effectiveness analysis. **Ethics**: DS conducted the literature review, identified and classified the relevant ethical issues, conducted the ethical analysis, and wrote the ethics section of this report.

Final report: All authors agreed with the final content of the final report.

4. Table of contents

5. Abbreviations

ACER	**Average cost-effectiveness ratio**
ALCA	Anti-Lung Cancer Association
AUD	Australian Dollar
BIA	Budget impact analysis
CAD	Canadian Dollar
CanSPUC	Cancer Screening Program in Urban China
CEA	Cost-effectiveness analysis
CENTRAL	Cochrane Central Register of Controlled Trials
CHEC	Consensus on Health Economic Criteria
CHF	Swiss francs
CHEERS	Consolidated Health Economic Evaluation Reporting Standards
CI	Confidence interval
CMS	Centers for Medicare & Medicaid Services
CO	Carbon Monoxide
COS	Consequences Of Screening
COS-LC	Consequences Of Screening Lung Cancer
CPD	Cigarettes smoked per day
CT	Computed tomography
CUA	Cost-utility analysis
CXR	Chest X-ray
DSA	Deterministic sensitivity analysis
e.g.	exempli gratia (= for example)
ELCAP	Early Lung Cancer Action Project
etc	et cetera
EUR	Euro
EQ-5D	European Quality of Life 5 Dimensions 3 Level Version
EQD	European Quality of Life Questionnaire
GBP	Great Britain Pound
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HADS	Hospital Anxiety and Depression Scale
HALY	Health-adjusted life-years

HPFS	Health Professionals Follow-up Study
HRQoL	Health Related Quality of Life
HTA	Health Technology Assessment
ICER	incremental cost-effectiveness ratio
ICTRP	International Clinical Trial Registry Platform
i.e.	id est (= that is)
IES	Impact of Event Scale
I-ELCAP	International Early Lung Cancer Action Program
ITT	Intention to Treat
IRR	Iranian rial
JECS	Japan Environment and Children's Study
JPY	Japanese Yen
IV	Inverse-variance
KCCR	Korean Central Cancer Registry
K-LUCAS	Korean Lung Cancer Screening Project
LDCT	Low-dose computed tomography
Lung-RADS	Lung Imaging Reporting and Data System
LUSI	Lung tumor screening and intervention trial
LYG	Life-year gained
MDT	Multidisciplinary Team
M-H	Mantel-Haenszel method
MISCAN	MIcrosimulation Screening ANalysis
NELSON	Dutch-Belgian Lung Cancer Screening trial (NEderlands Leuvens Longkanker Screenings Onderzoek)
NHS	UK National Health Services
NHS EED	Economic Evaluation Database from the UK National Health Service
NICER	National Institute for Cancer Epidemiology and Registration
NIHR	UK National Institute for Health Research
NLST	US National Lung Screening Trial
NR	Not reported
NZD	New Zealand dollar
OIS	Optimal information size
PanCan	Pan-Canadian Early Detection of Lung Cancer
PICO	Population, intervention, comparator and outcomes

PLCO	Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial
PSA	Probabilistic sensitivity analysis
PY	Pack-years
QALY	Quality-adjusted life year
RCT	Randomized controlled trial
RR	Relative risk
SD	Standard deviation
SEER	Surveillance Epidemiology and End Results
SF-12	Short Form questionnaire-12 items
SMD	Standardized mean difference
STAI-6	Six-item State-Trait Anxiety Inventory
TSCE	Two-stage clonal expansion model
UKLS	UK Lung Cancer Screening Trial
USD	United States Dollar
USPSTF	US Preventive Services Task Force
WHO	World Health Organization

6. Executive summary

Background

Lung cancer is the most important malignancy causing roughly 3,200 deaths in Switzerland each year and is most prevalent in smoking individuals. Individuals with a late-stage diagnosis of lung cancer have a poor prognosis. Low Density Computed Tomography (LDCT) may be a promising screening intervention for early diagnosis and treatment of lung cancer in high-risk populations to reduce morbidity and mortality due to lung cancer.

Aims

Based on the UK Health Technology Assessment (HTA) 'Low-dose computed tomography for lung screening in high-risk populations: a systematic review and economic evaluation' by Snowsill T et al. (issued in November 2018) an updated HTA report on the relative effectiveness and cost-effectiveness of LDCT screening for lung cancer in Switzerland was conducted which also addresses the ethical issues related to LDCT screening.

Methods

Clinical effectiveness

An updated literature search based on the one provided in the report by Snowsill was conducted. The search was adapted and extended for additional terms and comprised Medline via OvidSP, Embase, Web of Science via Clarivate Analytics, and the Cochrane Central Register of Controlled Trials (CENTRAL) via Cochrane Collaboration and the trial registries clintrials.gov and the WHO registry. Two assessors checked independently all literature items for randomised controlled trials comparing LDCT screening versus control or chest X-ray (CXR) for lung cancer in smoking individuals or heavy former smokers. Critical outcomes were lung cancer and overall mortality, and complications from invasive workup of false positive scans. Important outcomes were the number of false-positive scans, indeterminate scans, follow-up assessment and investigations with LDCT, the number of lung cancer detected and their stages, psychological distress, overdiagnosis, smoking cessation rate, type of cancer treatment, and quality of life.

Data abstraction of eligible trials was done in duplicate. Trials with ≥5 years of follow-up were considered for further assessment and critical binary outcomes that were available in both trial arms were pooled using a random effect model. Risk of bias was assessed with the GRADE tool. No continuous data was pooled, as too little data were reported in individual trials. In a sensitivity analysis for the critical outcome of lung cancer and overall mortality an indirect comparison of trials comparing

LDCT screening versus no screening, LDCT screening versus CXR screening, and CXR screening versus no screening was conducted.

Cost-effectiveness and budget impact

A systematic literature search of the economic literature on lung cancer screening based on the HTA published by Snowsill et al. in 2018 was conducted in Medline (via Ovid), EMBASE (via Ovid), and Web of Science (via Clarivate Analytics) in December 2020. Moreover, a non-systematic search update was conducted in Pubmed in October 2021 to identify potentially relevant articles published in 2021. All articles were screened by title, abstract, and if necessary, by full text review by two independent reviewers. Data extraction and quality assessment according to the Consensus on Health Economic Criteria (CHEC)-list for economic evaluations was conducted for all eligible articles. Population demographics, study characteristics, and main results were summarised and briefly described.

The cost-effectiveness analysis was based on a newly programmed version of the MIcrosimulation SCreening ANalysis (MISCAN) Lung model (a stochastic, microsimulation model). Like in our previous Swiss cost-effectiveness analysis based on NLST-effectiveness data, we modelled a cohort of 100,000 Swiss persons born between 1940 and 1980. Effectiveness data from the Dutch–Belgian lung cancer screening trial (NELSON) were used to calibrate the model. The inclusion criteria for patient eligibility to screening were based on the NLST, on the NELSON, and on the PLCOm2012 risk assessment criteria. Costs included costs for LDCT screen and invitation, risk-assessment, LDCT follow-up, biopsy, and treatment (divided by care phase and including immunotherapy costs as part of the terminal care costs). The analyses were conducted using a healthcare perspective, a lifetime horizon, and a discount rate of 3% (for both costs and effects).

The budget impact analysis was based on the results of the cost-effectiveness analysis. Undiscounted costs of selected screening scenarios were compared to no screening.

Ethics

Empirical research on patient attitudes as well as analytical literature on ethical issues was identified using purposive sampling on Pubmed and Google Scholar. Abstracts were selected for screening if they referred to ethical issues relating to screening or patient attitudes to screening, but only those that focused on these topics were included in the review. From the papers included, ethical issues were identified and categorised. Only one unanticipated ethical issue emerged from the literature review; one further ethical issue emerged during the ethical analysis. Following the identification of issues, they were categorised into two main groups: Clinical ethical issues concerning screening, and wider

issues concerning justice and discrimination. Each issue was subjected to normative analysis via the application of key ethical principles and the available arguments in the ethical literature.

Results

Clinical effectiveness

Thirteen trials comparing LDCT with no screening or CXR were identified and of those 7 trials had ≥5 years of follow-up which included 88,006 subjects for the primary critical outcome analyses. Three additional ongoing trials were found. For the network analysis 3 trials comparing CXR to no screening, 6 trials comparing LDCT with no screening and one trial comparing LDCT with CXR were available. For the critical mortality outcomes risk of bias in trials was judged as moderate. There was considerable variation in screening programmes in terms of screening intensity, with most trials conducting 3 to 5 screening rounds, the definition of a positive node and as a consequence the necessary work-up investigations. Only one trial (NELSON) used a volume-based and not diameter-based definition of a positive node.

The risk ratio (RR) of death from lung cancer of LDCT compared with no screening or CXR in 7 trials with ≥ 5 years of follow up was 0.80 (95%CI 0.72 to 0.88; test for heterogeneity I^2 = 0%). In the network analysis the league table for the pooled direct and indirect comparisons of trials comparing LDCT with CXR or no screening or CXR with no screening indicated that CXR compared to LDCT had a statistically significant higher risk ratio for death from lung cancer, LDCT compared to no screening a statistically significant lower RR of death from lung cancer and CXR compared to no screening had no effect on lung cancer mortality. The RR of death from all causes of LDCT compared with no screening or CXR (7 trials) was 0.96 (95%CI 0.92 to 1.00; I^2 = 0%). Two trials (NELSON and NLST) contributed roughly 75% of weight to the pooled summary of all mortality outcome data. In the network meta-analysis no statistically significant difference in overall mortality was found between any direct or indirect comparison. Obviously, more lung cancers were detected with LDCT and patients with LDCT compared to control were more likely to be diagnosed with lung cancers in earlier stages (I and II) (RR 2.69, 95% CI 1.94 to 3.74, I^2 = 80%, 7 trials). Three trials assessed psychological effects that may be associated with LDCT screening but only one trial (DLCST) evaluated the entire trial population. All trials had validity issues due to the relative subjectivity of outcomes assessments, lack of blinding, and loss to follow-up. No uniform picture in terms of psychological consequences from screening with LDCT can be drawn. In DLCST following the first and prior to the second screening round mean scores for anxiety were lower in the screening group, but likely not clinically relevant. During screening rounds, 2 – 5 participants in the control group experienced statistically significantly more negative psychosocial consequences in seven of nine health scales compared to the LDCT group.

Two trials evaluate smoking behaviour change in relation to lung cancer screening at the broadest study population level but did not show that LDCT screening was associated with higher quit rates when compared to control.

The definition of a positive node or findings in LDCT varied between trials and diagnostic work algorithms also differed. The range of any found thorax abnormality or protocol defined indeterminated scans during screening programs was wide and between 4.5% in MILD and 47.5% in the UKLS trial. The range of false positive scans (defined as the ratio of the [difference between recall scans/work-ups and screening detected lung cancers] and screened individuals) was also large between trials and varied between 1.2% in NELSON, 3.0% in DLCST, and 45.3% in the NLST trial. Trials with defined workup algorithms had considerably lower false positive rates. The rate of invasive procedures from false positive scans in individuals in need of a recall scan or diagnostic work-up ranged from 2.6% to 9.6%; data on complications from false positive LDCT was, however, very scarce. Rates of invasive procedures per screened individual varied between 0.5% and 11.4%.

Cost-effectiveness and budget impact

A total of 43 cost-effectiveness analyses were included in the systematic review. According to the CHEC checklist, the quality of reporting differed substantially between studies. The included studies showed high heterogeneity in the interventions (e.g., single, annual, biennial, triennial LDCT screening), comparators (no screening or CXR), the main source of effectiveness assumptions (e.g., NLST, NELSON, ELCAP, etc.), perspective (e.g., healthcare, payer, insurer, societal), and time horizon (from 1 year to lifetime). In general, a common theme in the study results was that LDCT screening is more costly and more effective than no screening or CXR (NB: studies based on NLST generally assumed that CXR was equal to no screening). In most cases, the incremental cost effectiveness ratios (ICERs) were below USD/EUR/GBP/NZD/CAD 100,000 per life year gained (LYG) or per quality-adjusted life-year (QALY) gained. Studies based on the recently published NELSON study seemed to lead to improved ICERs for LDCT screening if compared with studies based on NLST or other trials. Many studies emphasized that the screening strategy (e.g., inclusion criteria for lung cancer screening), the cost of LDCT scans, the effectiveness of screening (sensitivity and stage shift leading to lung cancer detection in early stages) and the incidence/prevalence of lung cancer are key factors affecting the cost-effectiveness of screening.

To compare the previously published analyses based on NLST effectiveness with the new estimations based on NELSON effectiveness, a total of 2,972 scenarios were modelled. The results showed that scenarios based on NELSON effectiveness led to more LYG if compared to the original scenarios based on NLST effectiveness. The average cost-effectiveness ratios (ACERs) comparing each scenario with no

screening for the models based on NELSON effectiveness led to ACERs ranging between CHF 14,452 to CHF 37,959 per QALY gained. The no screening scenario estimated the detection of 6,784 lung cancer cases and a total of 4,674 lung cancer deaths per 100,000 persons. The introduction of lung cancer screening led to a higher number of detected lung cancer cases and a lower number of cancer deaths. For the scenarios on the efficiency frontier, the number of detected lung cancer cases ranged between 6,799 (+15 cases per 100,000 persons compared to no screening) and 6,981, (+197 cases per 100,000 persons compared to no screening), while the number of lung cancer deaths would range between 4,471 (-4.3%) and 3,593 (-23.1%). In our previous study, the number of false positive screens per 100,000 persons (based on NLST effectiveness) were particularly high, ranging between 7,651 and 63,435. The new analyses based on NELSON false-positive rates showed a drastic decrease, with false positive screens ranging between 360 and 8,290 per 100,000 persons.

Depending on the screening scenario, the number of individuals needed to screen per LYG would range between 2 and 3 (i.e. you need to screen 2-3 persons at risk to gain one life-year), while the number of individuals needed to screen per death avoided would range between 21 and 41. The number of LDCT screens per lung cancer death avoided would range between 155 and 434 LDCT screens per LYG.

In the budget impact analysis, the total costs related to lung cancer treatment in Switzerland in the absence of screening were estimated to increase from CHF 474 million in 2023 to CHF 724 million in 2037. Compared to no screening, the budget impact of all screening scenario was higher. Over a period of 15 years, the total costs of lung cancer in the no screening scenario were estimated to reach CHF 9.4 billion, while the costs for three selected scenarios on the efficiency frontier ranged between CHF 10.2 billion and CHF 12.6 billion (i.e., +9% and +34% compared to no screening, respectively).

Ethics

Screening raises many ethical issues regarding access, stigmatisation, shared decision making and treatment modalities. These can all be addressed with careful design of screening campaigns and patient interaction, but particular care should be taken to avoid overstating the prospective benefits of screening. Perceptions of lung cancer as a “self-inflicted” disease are held by some citizens, but this view is not prevalent and screening is perceived positively by a majority. Screening also raises issues concerning just distribution of resources, with hundreds of patients needing to be screened to prevent one death from lung cancer and a high financial cost per averted death, and little impact on overall mortality. Implementation of screening would benefit those in lower socioeconomic groups and certain ethnic groups to a greater extent than other populations, but failure to implement screening would not amount to discrimination against these groups. Excluding other high-risk groups other than (ex-)smokers would also not be discriminatory given the differential balance of costs and benefits.

Conclusion

LDCT screening for lung cancer is associated with a reduced mortality from lung cancer but does not reduce overall mortality. Psychological consequences of screening (e.g. anxiety or depression) remain unclear and LDCT screenings does not seem to increase quit rates from smoking. False positive findings from LDCT remain a concern and important differences in false positive rates, repeated scans and invasive work-ups were found between trials. Volumed-based definitions of suspicious nodes, repeated scans and strict work-up protocols as applied in the large NELSON trial reduce false positive scans.

The great majority of the published cost-effectiveness analyses concluded that lung cancer screening may be a cost-effective intervention. Analyses based on data from the NELSON trial confirmed the positive results obtained in previous analyses based on the results of the NLST. The results of the cost-effectiveness analysis suggested that most lung cancer screening strategies may be cost-effective in Switzerland (assuming a threshold of CHF 100,000 per QALY gained). The cost-effectiveness and budget impact were highly dependent on screening intervals and smoking eligibility criteria. Although being more expensive than biennial and triennial screening strategies, annual screening showed the greatest potential reduction in lung cancer mortality and the highest increase of QALY gained.

Whether lung cancer screening represents a fair distribution of harms and burdens for the benefit conferred is a subjective judgment. Even if screening is deemed cost-effective in a financial sense, there is little impact on overall mortality and the number of patients needed to screen and the number of false positives incurred to prevent each lung cancer death may be too high to merit implementation. Whatever decision is ultimately made about screening, whether at the patient level or the health systems level, any values underlying that decision must be articulated clearly, along with the empirical evidence informing that decision.

7. Preamble

The Swiss Cancer Screening Committee, a national consortium of experts for cancer prevention commissioned this HTA report on low dose computed tomography (CT) screening in individuals at high risk of lung cancer (current and former smokers). During the scoping process, the purchaser and contractor agreed that this report should for the clinical effectiveness part be based on an update of the HTA Report 'Low-dose computed tomography for lung cancer screening in high-risk populations: a systematic review and economic evaluation' by T. Snowsill et al.[1] which was published in 2018. It was also agreed that the health economic model should be adapted to the Swiss context and supplemented by an ethical domain.

This report follows therefore in parts the structure of the HTA report by Snowsill et al.[1]. In particular, some tables were adapted from their report which are all labelled with a star (*).

8. Medical background

Each year, lung cancer is responsible for 1.6 to 1.8 Mio deaths worldwide [2] [3] and for about 3,200 deaths in Switzerland[4]. Over 20% of cancer-related deaths in Switzerland are caused by lung cancer[4]. Lung cancer growth remains usually undetected until later cancer stages compromising treatment options and success. Five-year survival in patients with advanced cancer stages is around 5%, whereas for early stages the five-year survival is up to 50%.[5] Screening for lung cancer in a high-risk population has therefore the potential to shift the detection to earlier cancer stages and treatment and to reduce cancer related mortality. There is evidence from one large randomized US National Lung Screening Trial (NLST, >53,000 participants randomized), demonstrating that screening with low-dose computed tomography (LDCT) compared to chest radiograph reduces lung cancer mortality. Ten-year follow-up results of the second largest lung cancer screening trial NELSON (>15,000 participants randomized) indicate an important reduction in lung cancer mortality with LDCT compared to no screening. Several smaller trials, but still in total with several thousand participants, have investigated the comparative effectiveness of LDCT versus no screening. Lung cancer screening programs using LDCT have been established in the US, UK and Poland. Several countries, including Switzerland, have not yet implemented such population-based programs as several questions like the burden for the work-up of positive or suspicious CT scans and costs of follow-up procedures in a real-world setting remain insufficiently addressed. For example, the NLST trial reported that a quarter of all LDCT scans were positive, thereof 96.4% were false-positives scans.[6] For these reasons a HTA report was commissioned by the Cancer Screening Committee, an expert panel for cancer screening recommendations to look into the comparative effectiveness and cost-effectiveness of routine LDCT lung cancer screening in smoking individuals to reduce morbidity and mortality from lung cancer in the Swiss health care setting.

1. Clinical effectiveness

1.1. Aim

The aims of the HTA report are

- to systematically assess the clinical effectiveness (benefit and harm) of lung cancer screening with LDCT compared to no screening or any other screening method relevant for the Swiss setting
- to assess the cost-effectiveness and potential budget impact of LDCT screening programs for lung cancer
- to address the ethical issues raised by LDCT screening.

1.2. Methods

1.2.1. Overview of the eligibility criteria

The overview of eligibility criteria (PICO-Question) used in the literature selection process is shown in **Fehler! Verweisquelle konnte nicht gefunden werden.** in the appendix.

PICO-Question	
Population	Smokers and former smokers (see section 1.2.2.1.)
Intervention	Low-dose computed tomography (see section 1.2.2.2.)
Comparator	No screening/ usual care and chest X-ray (see section 1.2.2.3.)
Outcomes	Critical and important patient-relevant outcomes (see section 1.2.2.4.)
Study design	Randomized and quasi-randomized controlled trials (see section 1.2.3.)
Languages	English, German, French (see section 1.2.4.)

Table 1 PICO-Question for the assessment of clinical effectiveness

1.2.2. Eligibility criteria

1.2.2.1. Population

Any asymptomatic adult population (≥18 years) at high risk of lung cancer due to smoking will be eligible.

1.2.2.2. Interventions

Any screening with LDCT irrespective of the number of screening rounds or screening intervals.

1.2.2.3. Comparators

No screening or usual care or chest X-ray. Screening with chest X-ray will be considered for two outcomes in a network meta-analysis (see section 7).

1.2.2.4. Outcomes

Critical outcomes:

- Lung cancer mortality (at least 5 years follow-up)
- All-cause mortality (at least 5 years follow-up)
- Number of false-positive scans with invasive procedures (e.g. fine-needle biopsy, bronchoscopy or surgery) --> A false-positive scan is defined as a positive scan result (leading to further testing or treatment) when lung cancer was absent. As the definitions of false-positive scans might vary between trials, the definition of false-positive scans will be extracted for each trial.
- Number of false-positive scans with complications --> A false-positive scan is defined as a positive scan result (leading to further testing or treatment) when lung cancer was absent. As the definitions of false-positive scans might vary between trials, the definition of false-positive scans will be extracted for each trial. As the definitions for complications might vary between trials, the definition for complications following invasive and non-invasive diagnostic procedures will be extracted for each trial.

Important outcomes:

- Number of false-positive scans --> A false-positive scan is defined as a positive scan result (leading to further testing or treatment) when lung cancer was absent. As the definitions of false-positive scans might vary between trials, the definition of false-positive scans will be extracted for each trial.
- Number of indeterminate scans --> An indeterminate scan is defined as a scan which does not allow the classification of the lung cancer as being present or absent. Indeterminate scans result in further testing. As the definitions of indeterminate scans might vary between trials, the definition of false-positive scans will be extracted for each trial.
- Number of follow-up assessments with LDCT
- Number of lung cancer detected
- Lung cancer stage --> not patient-relevant, however, early detection requires less severe therapeutic measures
- Interval lung cancer detection (after negative-screening result or undetermined-screening result without follow-up CT scan)
- Psychological distress (depression, anxiety, stress, other)

- Overdiagnosis
- Smoking cessation rate
- Number and type of lung cancer treatment
- Number of follow-up investigations (invasive and non-invasive)
- Quality of life

Further parameters or outcomes may be added during the assessment, especially if they are relevant to inform the health economic evaluation.

The relevant outcomes were classified according to GRADE as critical and important outcomes.[7-22] Critical outcomes would have a major impact on decision making and the quality of evidence available for these outcomes is the basis for judging the overall quality of the evidence for a clinical question.

1.2.3. Study design

Relevant study designs included randomized controlled trials (RCT) and quasi-RCTs (with assignment of treatment based on, e.g., alteration or date of birth). Although the latter methods for randomisation are deemed inadequate, these study types were considered because it can be assumed that individuals in such studies were prospectively assigned to the intervention or the comparator. [23] [24]

1.2.4. Languages

Trials published in English, French, and German were eligible for inclusion.

1.2.5. Literature search

An updated literature search based on the one provided in the report by Snowsill was conducted. The search was adapted and extended for additional terms (See Appendix and Table A 1). The literature search comprised Medline via OvidSP, Embase, Web of Science via Clarivate Analytics, and the Cochrane Central Register of Controlled Trials (CENTRAL) via Cochrane Collaboration.

The databases were searched on December 14th, 2020 from December 1st 2017 to December 1st 2020. The search strategy combined expanded search terms for ‘lung neoplasm’ and ‘tomography, X-Ray Computed’ with a search filter for randomised controlled trials (RCTs). Specifically, the best optimized RCT filter with regard to sensitivity and specificity, by Wong et al. [24] was used for the search in Medline, i.e. "Cochrane Highly Sensitive Search Strategy for identifying randomised trials in Medline: Sensitivity- and precision-maximizing version (2008 revision)“ filter combined with the search terms “random” and “randomised” were used. Details of search strategies used can be found in **Fehler! Verweisquelle konnte nicht gefunden werden.**. Conference proceedings or conference booklets were not searched. Moreover, trial registries of Clintrials.gov and WHO were systematically searched in November 2021 for ongoing trials.

Two reviewers independently screened titles/abstracts of records found in the literature search for potentially eligible studies after the removal of duplicate publications. Subsequently, two reviewers independently screened the full text articles of the potentially eligible studies in order to identify eligible RCTs. Discrepant screening results were discussed and resolved by consensus or by third party arbitration. Protocols of included RCTs were searched for within trial registries.

1.2.6. Decision on patient-relevant outcomes to be extracted

All patient-relevant outcomes were extracted and included in the assessment.

1.2.7. Data extraction

Data on study characteristics and patient-relevant outcomes (health outcomes) were extracted into a standardised form by one reviewer (AB, AG, ATH, YT) and checked by another reviewer (HCB). Discrepancies were resolved by discussion.

Information on patient recruitment time, maximum follow-up time, setting and country, eligibility criteria, and description of the screening interventions (including information accompanying smoking cessation programs) were extracted. General study population characteristics (age, sex, smoking behavior/status, etc.) and characteristics of the lung cancer-positive population (cancer stage, histologic type, etc.) were extracted. Radiation exposure was not extracted, but is listed in the HTA report and existing literature on radiation exposure is referenced.

Outcome data were extracted for the latest follow-up time-point. However, earlier time-points were extracted if drop-out rates for the later follow-up time-point were high (>30%) or unbalanced between arms (>5%).

Continuous outcome data was to be extracted as mean values for each intervention group at follow-up or, if not reported, as mean change from baseline.

For binary outcomes, the number of patients experiencing an event was extracted and analyzed, and not the number of events themselves. If only the number of events was available, this information was extracted and summarized in the relevant sections. Pooling of the number of events was only considered if consistently reported by all trials.

For missing information, study authors were not contacted.

1.2.8. Risk of bias and quality of evidence assessment

One reviewer (HCB) assessed the internal validity (risk of bias assessment) of each trial. This was checked by a second reviewer (AG). Discrepancies were resolved by discussion or third-party arbitration.

To assess the risk of bias of individual trials the following criteria were used.[25 26 27-41]

- adequate random sequence generation (selection bias)
- adequate concealment of treatment allocation (selection bias)
- adequate blinding of patients and healthcare providers (performance bias)
- adequate blinding of outcome assessors (detection bias)
- complete outcome data (attrition bias)
- reporting bias

Risk of bias for each of the aforementioned criteria was assessed as low, high or unclear in each trial. It was taken into consideration that blinding of outcome assessors is of less relevance for some outcomes (e.g. lung cancer or overall mortality) than for patient-reported outcomes. To judge the completeness of outcome data and the resulting risk of attrition bias, the following operationalizations were used:

- The risk of attrition bias was judged low if the proportion of patients with missing data was 0 - 10% in either study arm and comparable between the randomized treatment arms.
- The risk of attrition bias was also judged to be low if the proportion of patients with missing data was between 10-20% per arm, was comparable between the randomized treatment arms, and was being addressed using adequate methods. In case of continuous data, methods considered to be adequate are multiple imputation methods but not simple replacement methods like “last observation carried forward” or “baseline value carried forward”. In case of binary data adequate methods to address missing data are conservative assumptions about missing data; i.e. those patients with missing data in the control arm are treated in the analysis as if they had had beneficial outcome results.
- Missing data in the treatment arms were considered comparable if the difference between the intervention and control group were 5% or less.
- The risk of attrition bias was judged high if more than 20% of the data were missing irrespective of how the missing data were addressed in the analysis.

Reporting bias was judged to be low if all outcomes (relevant for the present review) described in the trial protocol (or trial registry) were reported in the results section of the publication. If the trial was not registered or no trial protocol was available, reporting bias was judged to be unclear.

The quality of the evidence was judged by one reviewer (HCB) and checked by another according to GRADE (Grading of Recommendations Assessment, Development and Evaluation) on the outcome level by considering all the available trials for the respective outcome. Discrepancies were resolved by

consensus or third-party arbitration. The following criteria were considered to judge the quality of the evidence: [26] [27-41]

Criteria for rating down the quality of evidence:

- risk of bias (internal validity)
- inconsistency
- indirectness
- imprecision
- publication bias

Criteria for rating up the quality of evidence:

- large magnitude of effect
- dose-response gradient
- all plausible confounders or other biases increase the confidence in the estimated effect

Imprecision refers to the confidence in the effect estimate. For binary outcomes, precision was judged to be adequate if the number of events was sufficient (rule of thumb >300 events).[32] If the sample size or the number of events was sufficiently large, the 95% CI of the effect estimate was examined. If the 95% CI was narrow enough not to include both the "no effect" line and a possible clinically relevant effect (also called minimal clinically important difference) precision was be judged as adequate [32]. For continuous outcomes, the precision was to be adequate if the optimal information size (OIS) was sufficient (simple sample size calculation to estimate whether the total number of included patients would be sufficient for an adequately powered RCT).

Using the GRADEpro GDT software[42] results of the judgment are presented in a summary of the findings table.

1.2.9. Data synthesis

Study characteristics and results of the eligible trials were presented per study in tables and descriptively summarised.

Clinically relevant and important outcomes (lung cancer mortality, all-cause mortality, and different cancer stages) where possible were summarised quantitatively in a meta-analysis using a random-effects model. Mantel-Haenszel weighting method [25] (MH) was used to pool the estimates for the included studies. As we had low event rates for some of the studies included in the meta-analysis but without zeros in both arms we used continuity correction of 0.5 in case of no events in one arm and

risk ratios are used to report the summary estimates. Effect estimates (summary and single for each trial) with the corresponding 95% confidence interval are presented in forest plots.

Heterogeneity of pooled effect estimates was estimated using I^2. Estimates of I^2 are interpreted under the guidance of the Cochrane Handbook. [26] Heterogeneity with an I^2 of 0% to 40% is considered low, 41% to 60% is considered moderate, and 61% to 100% is considered high. The interpretation of the observed I^2 value will depend on other measures for heterogeneity, namely Tau^2 (a Tau^2 value of 0.04, 0.09, and 0.16 represent low, moderate and high heterogeneity, respectively), the precision of the individual effect estimates of the included RCTs, and visual examination. [26 43]

In case of substantial heterogeneity, methodological and clinical factors that might explain the heterogeneity were to be explored in subgroup and sensitivity analyses.

1.2.10. Subgroup analyses

To assess possible variations of treatment effects the following subgroup analyses were defined prior to conducting the analyses:

- Internal validity (trial of high vs. low internal validity)
- Population characteristics (age groups, sex, number of cigarette package years)
- Population at risk (e.g. patient with smoking history vs. exposure to asbestos vs. family history of lung cancer)
- LDCT screening (single vs. multiple screening)
- Different definitions for positive CT scans (e.g. based on diameter of non-calcified nodules vs. definitions based on volume and volume-doubling time).

The sequence of the subgroup analyses listed above corresponds to the sequence in which the subgroup analyses will be performed depending on the available evidence.

Subgroup differences were to be assessed by interaction tests available within Review Manager 5.3 and according to the Cochrane Handbook. [26]

2. Network meta-analysis

A network meta-analysis was performed in addition to the direct comparison of LDCT screening with no screening. The random-effects network meta-analysis was performed for the clinically relevant outcome of lung cancer and overall mortality, and consists of three connected nodes (LDCT, chest X-ray, and usual care/no screening) (Figure). We used the frequentist random-effects method. Through the use of this network meta-analysis, external evidence from trials comparing chest X-ray with no

screening can be borrowed to assess the comparative effectiveness of no screening with LDCT and to compare the effectiveness of chest X-ray with LDCT. The relative effects of the compared screening strategies were reported as RRs with corresponding confidence intervals. [44-46]

Statistical analyses are performed using R packages "meta", and "netmeta". The confidence in the results of the network meta-analysis was to be assessed with CINeMA.[47]

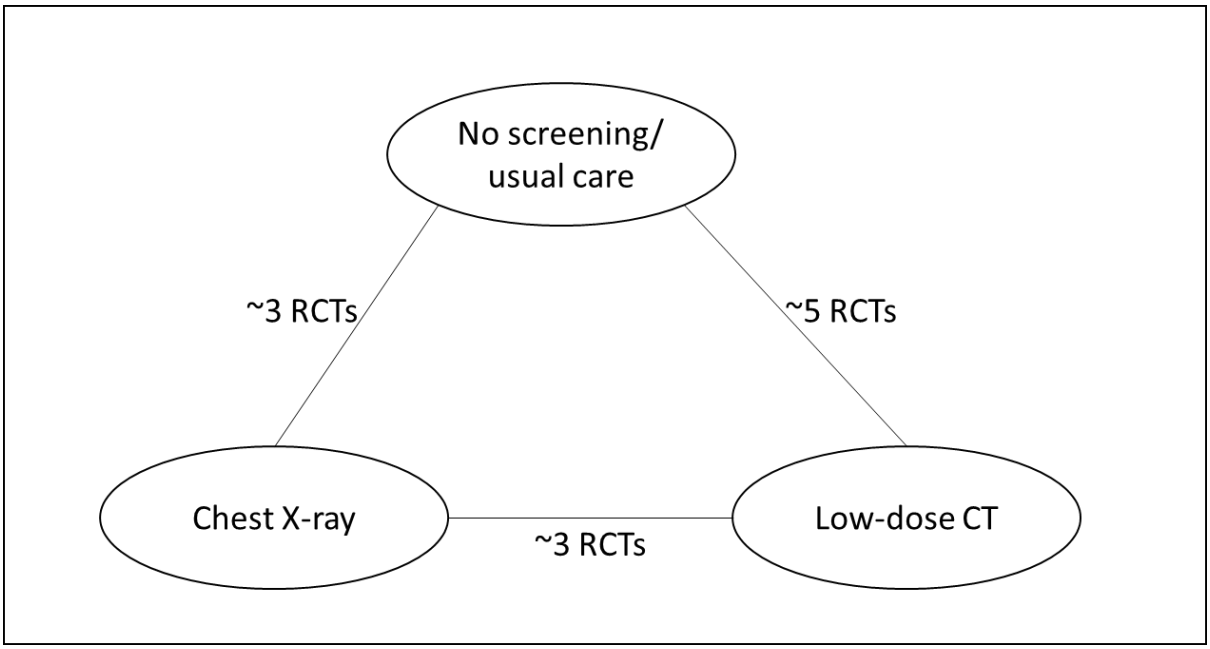


Figure 1 Network map of the network meta-analysis (based on all available evidence after literature screeening irrespective of trials with > 5 years of follow-up) at the time of the scoping process

8.1. Aim of the network meta-analysis

The network meta-analysis aims to estimate the relative efficacy between different interventions LDCT screening, CXR screening, and usual care with no screening of included trials with > 5 years of follow-up and was agreed with the purchaser of the HTA report during the scoping process.

8.2. Critical outcomes to be assessed in the network meta-analysis

The critical outcome to be assessed in the network meta-analysis was lung cancer mortality and overall mortality.

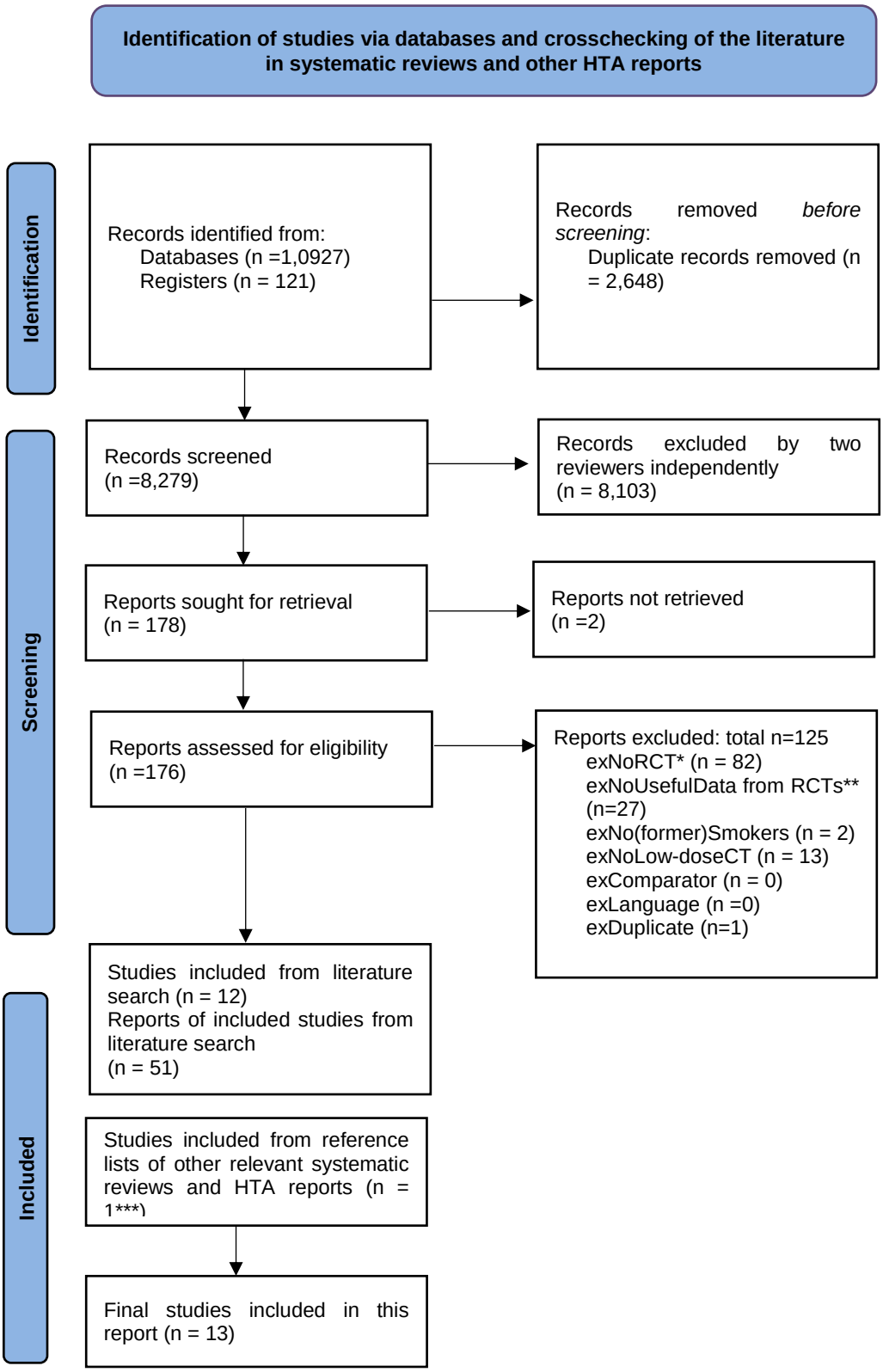
8.3. Sensitivity analyses – trial-specific (aggregated data) meta-analysis

In case of considerable heterogeneity (high I^2), and if too few RCTs were available for subgroup analysis, explorative sensitivity analyses were to be conducted. Sensitivity analyses might explain how specific parameters (e.g. population or screening characteristics) might cause heterogeneity. Further criteria for sensitivity analyses were to be defined a posteriori and will be strictly labeled as such.

9. Results

9.1. Literature search

Details on the number of identified records and trials from the literature search are provided in figure 1. For reporting, we followed the recommendations by Page et al. [48]. The detailed literature search is given in the appendix. In total 10,927 records from literature databases were identified and 8,279 records were screened and of those 176 records were assessed for eligibility and 13 studies were identified and included into the report. Many trials were reported in multiple publications that focused on different aspects of screening. A considerable overlap in reporting and in some trials inconsistencies in reporting was noted, mainly due to outcome reporting on multiple time points as trials evolved.



*e.g. systematic reviews or meta-analysis

**e.g. only protocols of otherwise relevant RCTs or secondary analyses of risk factors or outcomes irrelevant to the HTA report of otherwise relevant RCTs.

Figure 2 Trial selection process

9.2. Overview of included RCTs

We included 13 references encompassing 13 relevant RCTs which have been identified. We could not identify through our literature search the PLCO trial by Gohagan et al. [49] which was identified by Snowsill et al. Important baseline characteristics and types of comparator interventions (CXR or control) are listed in table 2. Nine trials were conducted in European countries and three trials were conducted in the USA and one in China. Two trials (including one pilot trial) were conducted in the UK. Only a minority of included trials provided information on outcomes that where considered clinically relevant or important outcomes.

An overview of additional baseline characteristics of included trials is given in Fehler! Verweisquelle konnte nicht gefunden werden..

Study identifier	Country	Recruitment time	Screening programme	Comparator	Sample size (n)	Age range, years (recruitment	Number of screening	Screening times and interval (years)	Duration of follow-up(mean/median)
DANTE[43]	Italy	2001 to 2006	LDCT, medical examination and one CXR	No screening, medical examination and one CXR	2,811 (2400 planned)	60–74	5	T0, T1, T2, T3, T4 (1-year interval)	At December 2012, median 6 years
Depiscan[50]	France	NR	LDCT	CXR	830	47–76 (protocol	3	T0, T1, T2 (1-year	NR
DLCS[51]	Denmark	2004 to 2006	LDCT	No screening	4,104	50–70	5	T0, T1, T2, T3, T4 (1-year interval)	Median: 9.47 years vs. 9.53 years
Garg et al.[52]	USA	2001	LDCT	No screening	190 (400 planned)	50–80	2	T0, T1 (1-year interval)	NR (planned 2 years)
ITALUNG[53]	Italy	NR	LDCT, smoking cessation programme	No screening, smoking cessation programme	3,206	55 –59	4	T0, T1, T2, T3 (1-year interval)	NR
LSS-PLCO[49]	USA	2000	LDCT	CXR	3,318 (3000 planned)	55–74	1	T0, T1 (1-year interval)	NR
LungSEARCH[54]	UK	2007 to 2011	Sputum surveillance, if abnormal sputum, LDCT	CXR at 5 years	1,568 (1300 planned)	Mean 63	5	T0, T1, T2, T3, T4 (1-year interval)	NR (planned 5 years)
LUSI[55, 56]	Germany	2007 to 2011	LDCT, smoking counselling	No screening, smoking counselling	4,052 (4000 planned)	50–69	5	T0, T1, T2, T3, T4 (1-year interval)	NR
MILD[57, 58]	Italy	2005 to 2011	LDCT (annual and biannual), smoking cessation, pulmonary function test, blood	No screening, smoking cessation, pulmonary function test, blood sample	4,099 (10,000 planned)	> 49	10	T0, T1, T2, T3, T4, T5, T6, T7, T8, T9, (1-year interval) vs. T0, T2, T4, T6, T8 (2-year	Median 6.2 years
NELSON [59,60 61]	The Netherlands/ Belgium	2003 to 2006	LDCT	No screening	15,822	50–75	4	T0, T1, T2, T3, T0 to T1, 1 year; T0 to T2, 3 years; T0 to T3, 5.5 years	10 years

NLST [6,62]	USA	2002 to 2004	LDCT	CXR	53,454	55–74	3	T0, T1, T2 (1-year interval)	Median 6.5 years
UKLS[63]	UK	2011 to 2012	LDCT	No screening	4,061 (4000 planned)	50–75	1	T0	7.3 years
Yang [9]	China	2013 to 2014	LDCT	No screening	6,657	45-70	2	T0, T1 (2 years)	NR

Table 2 Baseline characteristics and type of control interventions of randomised controlled trials of LDCT screening versus no screening or CXR

10.Definition of high risk populations for lung cancer in included trials

Characteristics of included populations varied between trials as definitions of high risk populations for lung cancer were not uniform. Such differences in the definition of high risk populations can determine differences in prevalence and incidence of lung cancer between trials.

The Detection and Screening of Early Lung Cancer with Novel Imaging Technology and Molecular Essays (DANTE) trial[43] defined high-risk based on age (aged 60–74 years) and smoking of at least 20 pack-years for smokers and former smokers who had to have quit for < 10 years before recruitment. The Italian lung cancer screening trial (ITALUNG) used similar criteria for the definition of high risk for lung cancer.[64] The Despiscan trial [50] defined high-risk based on age (aged 50–75 years) and current or former smoking who had to have quit < 15 years from enrolment. The Danish Lung Cancer Screening Trial (DLCST)[51] defined high-risk based on age, (aged 50–70 years) and current or previous smokers (who had to have quit after the age of 50 years and < 10 years prior to the start of the study) both with ≥ 20 pack-years. Patients had to be able to climb 36 steps without interruption and had to have a forced expiratory volume in 1 second (FEV1) of at least 30% of predicted normal at baseline. The MILD trial defined high risk based on age (≥ 49 years) and current or former smokers (having quit within 10 years of recruitment) with ≥ 20 pack-years of smoking[57], [58]. The National Lung Screening Trial (NLST)[62] defined high risk based on age (55–74 years) and a smoking history of ≥ 30 pack-year for both smokers and former smokers (who had to have quit smoking for at least 15 years). The UKLS [63] recruited individuals based on a prediction rule (Liverpool Lung Project lung cancer risk prediction algorithm) with a predicted lung cancer risk of ≥ 5% for developing lung cancer in the next 5 years that was based on age, sex prior diagnosis of pneumonia, family history of lung cancer with or without early onset (<60 years of age), years of smoking, prior history of malignances and non-malignant lung disease. The NELSON trial [60] defined high risk based on age (aged 50–75 years) and current and former smoking (having quit ≤ 10 years), individuals who smoked > 15 cigarettes per day for > 25 years or > 10 cigarettes per day for > 30 years. Additional information on further baseline characteristics of trials populations in regard to current smoking status, age, gender and family history of lung cancer is provided in Tables 2 and 3 in the appendix.

10.1. Characteristics of screening programmes

Ten studies compared LDCT screening with usual care (no screening), and three studies (DEPISCAN [50], LLS-PLCO [49], NLST[6]) compared LDCT screening with CXR screening. All trials with the exception of DEPISCAN [43], Garg [52], UKLS [63]and Yang [9] had beside a baseline screening at least 3 screening rounds, whereas the latter trials had only a baseline or one additional screening round.

Definitions of a positive scan varied widely across studies in terms of nodule sizes or volume and likewise did work-up algorithms for non-normal CT scans vary substantially (Table 3). Most trials used diameter size of non-calcified nodules with or without morphological suspicious malignancy aspect as a criteria for a positive finding. Critical diameter size again varied substantially between trials. Only two trials (LUSI [55]; NELSON [59,60]) used volume parameters and volume doubling time as a criteria for a positive screening finding, which may have lead to lower false positive rates in these particular trials. Several trials also defined other abnormal findings as suspicious lung cancers such as adenopathies or effusions.

There was also variations in imaging evaluation and interpretation. Most trials used double reading techniques by experienced, independently rating radiologists. The diagnostic follow-up strategies for suspicious abnormality findings varied also between trials. Some trials had very detailed algorithms (DANTE [43], DLCST [51], ITALUNG [53], LUSI [55], NELSON [59,60], UKLS[63]) other trials had patients with suspicious findings referred to their care givers with no standardized work-up algorithms in place (Lung SEARCH [54], NLST [6,62] and Yang [9]). Most studies used further diagnostic imaging (e.g. high-resolution CT or chest fludeoxyglucose (18F) positron emission tomography ([18F] FDG-PET)) and/or invasive biopsy with rapid on-site examination, but some trials did not detail diagnostic work-ups.

Methods and settings for recruitment also varied between trials. Further information in addition to details on screening adherence is provided in the appendix in Table A 4.

Study identifier (country)	Screening programme comparison	Definition of a positive scan for lung cancer	Imaging evaluation and interpretation strategy	Diagnostic follow-up for suspicious abnormality finding
DANTE (Italy) [43]	LDCT (and baseline CXR and sputum cytology testing) vs. No screening (and baseline CXR and sputum cytology testing)	LDCT Non-calcified pulmonary nodule (≥ 10-mm diameter), non-nodular lesions suggestive of malignancy, such as hilar masses, focal glass opacities, major atelectasis, endobronchial lesions, mediastinal adenopathy, pleural effusion or mass Baseline CXR Non-calcified shadow, or hilar mass, or enlarged mediastinum, or pleural effusion/thickening or lytic bone lesion	LDCT Whole lungs scanned at full inspiration (following single breath hold) Independent double-reading of images by experienced chest radiologists. Decision based on consensus Baseline CXR Read by radiologists who were blind to the CT scan results	- Smooth lesion <10 mm: LDCT at 3, 6 and 12 months; if no changes, follow-up after one year - Non-smooth lesion <6 mm: follow-up by LDCT at 3, 6 and 12 months, if no change, follow-up after one year. - Non-smooth lesion >6 mm but < 10 mm: oral antibiotics and HRCT after 6 to 8 weeks. If no regression, evaluation on a case-by case basis invasive procedures (bronchoscopy, percutaneous fine-needle or core biopsy, or VATS) - Lesion > 10 mm but < 20 mm: oral antibiotics and HRCT after 6 to 8 weeks. If no regression, PET-scan. If the PET positive, a tissue diagnosis is sought. If the PET-scan is negative, close follow-up - Lesion > 20 mm: discretional oral antibiotics and HRCT or standard contrast-enhanced CT, and PET-scan. If PET positive, tissue diagnosis. If PET-scan negative, close follow-up - Focal ground glass opacities: oral antibiotics and HRCT after 6 to 8 weeks. Evaluation on a case-by case basis based in size, number and location of lesion
Depiscan (France) [50]	LDCT vs. CXR	LDCT Positive (requiring follow-up) if non-calcified nodules evident CXR Positive (requiring follow-up) if non-calcified nodules evident	LDCT Whole lungs scanned at full inspiration (following single breath-hold)	LDCT Follow-up protocol (recommended) - Nodule of ≤ 5 mm – LDCT at 1 year

			Independent double-reading of images by radiologists. Decision based on consensus CXR Not reported	- Nodule of > 5 mm but < 10 mm – LDCT at 3 months; if stable, LDCT at 6, 12 and 24 months. If enlargement histological diagnosis - Nodule of ≥ 10 mm – contrast enhanced CT, PET scan, and/or histological diagnosis. Results discussed by multidisciplinary team of pulmonary oncologist, radiologist and thoracic surgeon CXR Follow-up protocol - Suspected non-calcified nodule – LDCT scan, following the same recommended protocol as in the LDCT arm
DLCST (Denmark)[51]	LDCT (and PFT and < 5-minute cessation counselling) vs. No screening (and PFT and < 5 minute cessation counselling)	LDCT Category 1) Nodules ≤ 15 mm in maximal diameter with benign characteristics (for calcified nodules up to 20 mm) no further action Category 2) nodules below 5 mm no further action. Category 3) Nodules with diameter ≥5 and ≤15 mm no benign aspect rescan after 3 months (category 3). Category 4) Nodules exceeding 15 mm diagnostic investigation Category 5) growing nodules diagnostic investigation, in addition to nodules with suspicious morphology. After repeat CT scan, nodules were described as regressed, stable or growing by the radiologists. Growth was defined as an increase in volume of at least 25%	LDCT Scan at full inspiration Read by two experienced chest radiologists. Decision based on consensus	LDCT Follow-up protocol implemented after referral (decided by pulmonologist and radiologist) - Indeterminate LDCT screen – LDCT at 3 months, often followed by PET-CT - Positive LDCT screen – CT with contrast, followed by individual plan (VATS in most cases)

Garg *et al.* (USA)[52]	LDCT vs. No screening	LDCT Positive (requiring follow-up) if between one and six non-calcified nodules evident	LDCT Scan of lungs and diaphragm at full inspiration Read by one experienced chest radiologist; some systematically selected scans read by a second radiologist	LDCT Follow-up protocol - Positive LDCT screen – thin-section CT for diagnostic purposes
ITALUNG (Italy)[53]	LDCT (and invitation to smoking prevention programme) vs. No screening (and invitation to smoking prevention programme)	LDCT Non-calcified nodule ≥ 5-mm diameter, or non-solid nodule ≥ 10-mm, or part-solid nodule, nodules increasing by ≥ 1-mm mean diameter, increase in solid part of a nodule from one scan to the next, several nodules indicative of inflammatory disease	LDCT Independent double-reading of images by experienced radiologists. Decision based on consensus	LDCT Follow-up protocol at each centre for positive LDCT scans - No nodule growth (or regression) – LDCT at 1 year - Solid nodules of ≥ 8 mm and ≤ 10 mm – FDG-PET, followed by FNAB if positive or LDCT at 3 months if negative. If FNAB not positive then LDCT at 3 months - Non-calcified nodules of ≥ 5 mm and ≤ 7 mm (solid or part solid) – LDCT at 3 months - Growing nodules (peripheral) – FDG-PET or CT-guided FNAB - Growing nodules (deep) – FDG-PET or FBS - Airway abnormalities – sometimes followed up using FBS - Nodules indicative of inflammatory disease – antibiotics followed by LDCT at 1 month - Partial resolution – LDCT at 2 months If FNAB indicated lung cancer, a staging CT scan was performed

LSS-PLCO (USA)[49]	LDCT vs. CXR	LDCT At T0: Non-calcified nodule ≥ 4 mm diameter, any other abnormality considered suspicious by radiologist After T0 criteria changed: Non-calcified nodule ≥ 4 mm diameter, or spiculated nodule ≤ 3 mm diameter, or focal parenchymal opacities, or endobronchial lesions, or other abnormality considered suspicious by radiologist CXR Any nodule or mass, infiltrate/consolidation, alveolar opacity, enlargement of hilar or mediastinal lymph nodes (not calcified), lung/lobe collapse or closure	LDCT Read by one radiologist; some scans also independently read by a second radiologist CXR Single, postero-anterior view CXR	LDCT Individuals with positive screening results referred to personal care provider for further evaluation, no no specific follow-up work-up algorithm in place CXR As with LDCT
LungSEARCH (UK)[54]	LDCT (and sputum surveillance and AFB) vs. No screening (and exit CXR at 5 years)	LDCT Positive if ≥ 9-mm abnormal nodule	LDCT Not reported	LDCT If AFB positive/abnormal withe cells exhibiting squamous metaplasia, mild to severe dysplasia, carcinoma in situ or carcinoma then LDCT. A positive/abnormal LDCT (nodule size ⩾9 mm) could initiate cancer investigations according to local practice. Individuals with both normal AFB and LDCT continued to have these tests annually
LUSI (Germany)[55]	LDCT (and cessation counselling) vs. No screening	LDCT Nodules ≥ 5 mm, or Volume Doubling Time (VDT) VDT ≤ 600 days	LDCT Read by radiologists, with special training for the study	LDCT Follow-up protocol for positive LDCT results - Nodules of ≥ 5 mm and ≤ 7 mm – LDCT at 6 months if VDT > LDCT at 12 months

	(and cessation counselling)			- Nodules of ≥ 8 mm and ≤ 10 mm – LDCT at 3 months, if VDT 400 – 600 recall 6 months, if diameter < 7.5 mm 2 months recall, if diameter ≥ 7.5-10 mm, ≤400 VDT or diameter >10 mm immediate recall - Nodules of > 10 mm diameter or malignancy aspect immediate VDT of ≤ 400 days – antibiotics followed by CT, PET or immediate biopsy, as decided by pulmonologist
MILD (Italy) [57, 58]	LDCT – annual (and cessation programme, pulmonary function test and blood sample) vs. LDCT – biennial (and cessation programme, pulmonary function test and blood sample) vs. No screening (and cessation programme, pulmonary function	LDCT Nodules of ≥ 60 mm^3, i.e. approximately 5 mm diameter and nodules of ≥ 60 mm^3 but ≤ 250 mm^3 indeterminate, about 5-8 mm diameter), hilar/mediastinal lymphadenomegaly (non-calcified), atelectasis, consolidation, other indicative pleural findings, nodules increasing in volume by 25% in 3 months	LDCT Whole lungs scan (following single breath hold) Independent double-reading of images by trained radiologists (third radiologist arbitrated disagreements)	LDCT Follow-up protocol for positive LDCT results - Volume of ≥ 60 mm^3 but ≤ 250 mm^3 (indeterminate) – LDCT at 3 months - All positive results (not indeterminate) – follow-up PET-CT, contrast-enhanced CT or biopsy

	test and blood sample)			
NELSON (the Netherlands and Belgium) [59,60]	LDCT vs. No screening	LDCT Non-calcified nodule with a volume ≥ 500 mm^3, or percentage volume change ≥ 25% combined with VDT < 400 days. New solid component in a previously non-solid nodule. Indeterminate results: Nodules with volume 50–500 mm3, or a percentage volume change ≥ 25% combined with a VDT 400–600 days. 20	LDCT Scan from posterior recess to apex of the lung, no use of contrast Independent double-reading of images by experienced radiologists, except the last two rounds (read by a single, experienced radiologist)	LDCT Follow-up protocol for positive LDCT results - Indeterminate results – LDCT at 6-8 weeks and at 3–4 months depending on screening rounds The follow-up result in the second step was classified as negative or positive on the basis of nodule volume doubling time - Positive results – follow-up and diagnosis by pulmonologist, by standardised protocol with physical examination, contrast-enhanced CT, FDG-PET, bronchoscopy
NLST (USA)[6,62]	LDCT vs. CXR	LDCT Non-calcified nodule ≥ 4 mm diameter, other abnormalities could be classified as positive or suspicious or adenopathies or effusion. CXR Any non-calcified nodule or mass At T2, abnormalities suspicious for lung cancer that were stable across the three rounds were classified as minor abnormalities and not as positive results.	LDCT Read by experienced radiologists, images also compared with previous LDCT screens CXR Single-view posteroanterior X-rays, read by experienced radiologists	LDCT Follow-up protocol for positive LDCT results (guidelines only, details not reported) by patient's health care provider or according to radiologists' recommendations for diagnostic follow-up
UKLS (UK)[63]	LDCT vs.	LDCT - Category 1 Benign nodules fulfilling one of the following criteria: a benign pattern of	LDCT Scan (with single breath hold)	LDCT Follow-up protocol for LDCT results - Category 1 nodules No nodules: No further action required.

	No screening	calcification, fat, measuring < 3mm in diameter or volume of < 15mm3; or intrapulmonary lymph nodes fulfilling the following criteria: they lie within 5mm of the pleura, are < 8mm in diameter, smooth bordered and ovoid, and have at least one interlobular septum radiating from their surface - Category 2) Solid and intraparenchymal nodule, maximal diameter of 3.1–4.9mm or a volume of 15–49mm^3. Solid and pleural or juxtapleural, maximal diameter of 3.1–4.9 mm. If non-solid or part solid, a maximal diameter of 3.1–4.9 mm. Solid component diameter of < 3mm and/or volume of < 15mm^3. Non-solid/ground glass opacities, independent of diameter - Category 3) Solid and intraparenchymal, volume of 50–500mm^3. If solid and pleural or juxtapleural, diameter 5–9.9 mm. If non-solid or part solid, diameter of the ground-glass component of > 5 mm. If part solid, solid component volume of 15–500mm^3 or maximal diameter of 3.0–9.9 mm - Category 4) If solid and intraparenchymal, volume of > 500mm^3. If solid and pleural or juxtapleural, diameter of ≥10 mm. If part solid, solid component diameter of ≥10mm or volume of > 500mm^3	Double-reading by experienced chest radiologists. Decision based on consensus	- Category 2 nodules Follow-up CT scan at 12 months. - Category 3 nodules Follow-up CT scan at 3 months. - Category 4 nodules Referral to MDT.

Yang China [9]	LDCT vs. No screening / questionnaire inquiry	LDCT - Non calcified nodule or mass ≥4mm diameter	LDCT Scan Double read by two senior radiologists & three experienced clinicians	LDCT - Management according to NCCN guidelines in oncology
AFB, autofluorescence bronchoscopy; FBS, optical fibrobrochoscopy; FNAB, fine-needle aspiration biopsy; PFT, pulmonary function test; VATS, video-assisted thoracic surgery; VDT, volume doubling time.				

Table 3 Definition of positive nodes in trials comparing LDCT with no screening or CXR

10.2. Computed tomography parameters

The technical specifications of lung scanners in included trials are presented in table A 5 in the appendix. The DANTE,[43] Garg et al.,[52] ITALUNG[64] and German Lung Cancer Screening Intervention (LUSI)[55] trials used single-slice technology. Trials initiated later used multi-slice helical CT technology where the source and detector travel along a helical path relative to the object, i.e. the patient is moved on the CT table through the bore of the scanner meanwhile the gantry rotates. Helical CT allows for shorter examination time and larger scanning volume, and provides better z-axis resolution at a given x-ray dose. Quicker scanning has the additional advantage of less breathing artefacts due to faster processing time and better quantification of thoracic lesions where present. The slice thickness ranged in included trials between 1 and 3 mm which are generally considered as thin and superior to thicker slices. The DANTE trial [43], however, used a slice thickness of 5 mm. Voltage (in kiloVoltage, kV) and tube current time product (in milliAmpere seconds) reflect the rate of X-ray production and number of X-rays produced. In some trials, automated body mass index adapted parameters were used. A helical CT beam trajectory is characterized by its pitch, which is equal to the table feed distance along the scan range over one gantry rotation divided by the section collimation. [65] When pitch is greater than 1, the radiation dose for a given axial field of view (FOV) is decreased compared to conventional CT. At high pitches there is, however, a trade-off in terms of noise and longitudinal resolution [66]. In all but one trial the pitch was above 1. Estimates of radiation doses were only provided by very few trials.

10.3. Ongoing trials

We rescreened trial registries of clintrials.gov on November 3rd, 2021 with the search terms 'low dose computed tomography lung cancer' and also searched for 'lung neoplasm', 'neoplasm' and 'CT scan' and found 75 and 46 hits, respectively. Besides the trials which are already included in this review, we identified two additional ongoing trials (Table 4). Trial NCT02898441 is an ongoing trial from Shanghai China with baseline screening results published and included into this report. The trial is expected to be terminated in 2023. Trial NCT03975504 is an extension of only the screening arm from the previous arm and thus will not be of future relevance. The Yorkshire Lung Screening Trial (UK) started enrolling patients in September 2018 and plans to enroll over 6,800 participants who will undergo baseline and two year LDCT screening. Individuals are enrolled at registered practices in the Leeds area. The trial will use a Zelen design and will passively follow control group patients.

We additionally identified one study protocol from 2012 for a planned trial [67] from Japan. An update on this trial was published in 2017 [68] confirming that the trials was ongoing. In total, 27,000 participants 50–70 years of age with a smoking history under 30 pack-years were randomly assigned to low-dose thoracic CT done in the first and sixth year, whereas the control group was invited to have chest X-ray

done in the first year. The participants in both groups were also encouraged to undertake routine lung cancer screening using chest X-ray annually. A 10-year follow-up period is planned, with lung cancer mortality being the primary outcome.

Registration number **Country**	**Planned start** **Completion date**	**Participants to be enrolled**	**Published results**
NCT02898441 China[9].	January 1, 2014 December 2019 (primary completion) Study completion July 31st, 2023	6,000 participants with high-risk for lung cancer	Screening observation after first round, mortality rates for 2 year follow-up
NCT03975504 China	August 1, 2018 July 31, 2023	3,000 subjects of the screening arm	No
ISRCTN42704678 and NCT03750110 Yorkshire Lung Screening trial UK [69]	September 1st 2018 September 2021 (primary completion) September 2022 Study completion July 1st 2024	6,892 participants at high risk of lung cancer registered at participating general practices in the Leeds area	2 rounds of LDCT screening at an interval of 2 years. 4 to 6 years follow-up for screened individuals (use of Zelen design)
NR JECS Study [67] Japan	Protocol May 2012 recruitment start unclear	27,000 participants	Screening with LCDT at 1 and 6 years Controls annual CXR
NR not reported			

Table 4 Ongoing trials pf LDCT screening versus no screening or CXR

10.4. Outcome measures in included trials

Reporting of outcomes that were considered during the scoping process of this HTA report as critical or important varied greatly between trials. The majority of trials did not report on psychological outcomes like quality of life, stress or anxiety. According to the scoping process for this report, critical outcomes of lung cancer and overall mortality are assessed only in trials with sufficient follow-up time of ≥5 years. In addition, risk of bias assessment was restricted to critical and important outcomes that were available while maintaining the randomized design and reported in both, the screening and the control arms. All outcomes that were only available in the screening arm (e.g. all outcomes on LDCT screening test performance like false positive test rates of LDCT) are formally single trial arm

evaluations and subject to different forms of biases, typically encountered in cohort studies (for example lack of blinding for endpoint assessment) and were therefore not subject to bias assessments. *Lung cancer and all-cause mortality* in trials with ≥5 years follow-up were reported in seven trials. [43], [51], [57], [6], [59], [60],[55], [70] Details on the number of events for these critical outcomes can be found in the appendix in Table A 7.

Data on cancer incidence and tumor stage was published in 11 trials [71], [50], [7], [72], [73], [49], [74] [58],[75], [62],[9], [70] and of those 7 trials provided follow-up data of ≥5 years. [71], [50], [56], [58], [75], [62], [70]

Smoking cessation during follow-up was reported in three trials with ≥5 years of follow-up (DLCST [51], LLS PLCO [49] and UKLS [63]). The ITALUNG trial [53] also provided data on smoking cessation.

Four trials reported outcomes on *psychological consequences of screening*. Only one trial (DLCST [76]) assessed health related quality of life in the entire trial population and used a de novo developed and validated instrument the lung-cancer-specific questionnaire (COS-LC). The NELSON trial[77] used the European Quality of life Questionnaire (EQD) and the Spielberger Anxiety and Distress scales. The UKLS trial reported depression anxiety scores using the HADS instrument. [78] The NELSON and the UKLS trials [78] applied these instruments in a sub-population of the entire trial population. The NLST [79] investigated the psychological consequences of LDCT screening exclusively in 2,812 participants from 16 cites who had a positive finding in a scan of the lungs during the T0, T1, or T2 screenings. Later, additional participants were recruited with incidental findings on scans other than abnormalities of the lungs. These individuals were matched with a negative screen control. Due to selection criteria of NLST with focus on individuals with positive LDCT findings the trial was excluded for this endpoint due to limited comparability with other trials and because the used selection criteria by NLST investigators did not allow for an overall assessment of the psychological consequence of LDCT screening in the entire trial population or in a representative subsample of the trial population. Due to the difference in used quality of life measurement scales and applications in trial sub-populations it was decided not to formally combine outcomes of the psychological consequences of screening.

10.5. Risk of bias of included studies

In all the trials risk of bias was assessed with the Cochrane Risk of Bias tool.[80] Not all trials did report on critical and important outcomes as can be seen in an overview in table 5. For all outcomes on screening performance of LDCT information was only available for the screening arm. These outcomes were not assessed for the risk of bias. Risk of bias assessment was therefore confined to outcomes that were assessed in the entire trial population (and not in trial subgroups) and in both arms e.g.the critical outcomes lung cancer and overall mortality, and the important outcomes lung cancer stages.

	Critical outcomes		Important outcomes		Important outcomes (not assessed as only reported in one trial (HRQoL) or only reported in trial subpopulations (smoking cessation)	
Study identifier	**Mortality Lung cancer**	**All-cause**	**Cancer incidence**	**Stage distribution**	**HRQoL**	**Smoking cessation**
DANTE [43],[81]	≥ 5 years	≥ 5 years	≥ 5 years	≥ 5 years	NR	≥ 5 years
Depiscan [50]	NR	NR	NR	NR	NR	NR
DLCST[51]	≥ 5 years, 1°, 10 years	≥ 5 years	≥ 5 years, 2°, 5 years	≥ 5 years, 2°, 5 years	COS-LC 1–5 years	Annual smoking status 1–5 years
Garg et al. [52]	NR	NR	NR	NR	NR	NR
ITALUNG [53]	NR, 1°, 8 years	NR, 2°, 8 years	NR, 2°, 8 years	NR	NR	NR
LSS-PLCO [49]	NR	NR	NR	NR	NR	NR
LungSEARCH [54]	NR, 1°, 15 years	NR	NR	NR, 2°, 5 years NR	NR	NR
LUSI [55], [56]	≥ 5 years	≥ 5 years	≥ 5 years	≥ 5 years	NR	NR
MILD [57], [58]	≥ 5 years, 1°,10 years	≥ 5 years	≥ 5 years	NR	NR	NR, 2°, 10 years
NELSON [59], [60]	≥ 5 years	≥ 5 years	≥ 5 years	≥ 5 years	< 5 yearsc 2°	< 5 yearsc
NLST [6], [62], [79]	≥ 5 years, 1°	≥ 5 years, 2°	≥ 5 years, 2°	≥ 5 years	< 1 years	NR
UKLS [70]	≥ 5 years	≥ 5 years	≥ 5 years	≥ 5 years	< 5 years	NR

1°, primary outcome; 2°, secondary outcome; COS-LC, consequences of screening lung cancer; LSS-PLCO, Lung Screening Study as part of the Prostate, Lung, Colorectal and Ovarian cancer screening trial; NR, not reported.

Table 5* Reported outcomes measured in screening and control arms in trials with and without ≥5 years of follow-up

10.6. Risk of bias for lung cancer and overall mortality

Seven trials with ≥5 years of follow-up (DANTE, DLCST, LUSI, MILD, NELSON, NLST and UKLS) provided data on lung cancer and overall mortality (Details are provided in the appendix in Table A 6). Generally risk of bias for lung cancer and overall mortality outcomes was assessed as low. Generation of randomisation sequences was rated adequate. However, in the MILD trial[57], [58] randomisation did not achieve balanced study groups and there were differences for several study participants' characteristics at baseline (participants' sex, current smoking status and FEV1). Allocation concealment was unclear in all trials with the exception of the UKLS trial [70]. In all trials targeted sample size according to the power and sample size calculation was achieved. By nature of the open trial design blinding of study participants and staff was not possible. All trials had blinded outcome assessment committees in place and attrition bias was due to small losses to follow from individuals with unknown vital status or cause of death low in all trials. In all trials mortality outcomes were pre-specified.

10.7. Risk of bias for psychological consequences and health-related quality of life

Three included trials (DLCST [51], NELSON [59], [60] and UKLS [63] [70]) provided data on psychological consequences and HRQoL in relation to LDCT screening and assessments for the risk of bias are provided in the appendix in Table A 6. Risk of bias for psychological consequences and health related quality of life was higher in all trials compared to mortality outcomes. Risk of bias for random sequence generation was considered low in DLCST [51] and UKLS [70]. Information on allocation concealment was missing for DLCST [51] and NELSON [59], [60]. Due to the open design, risk of bias for blinding was high in all trials. Risk of bias was also considered high for blinded outcome assessment as outcomes are more subjective. Loss to follow-up was high in all trials and the risk for attrition bias was rated as high.

10.8. Risk of bias for smoking behaviour

Two trials (DLCST [81] UKLS [82]) could be evaluated in regard to smoking cessation in relation to LDCT screening which included the representative entire population of individuals who were smoking at baseline. In both trials risk of bias was rated higher in comparison to mortality outcomes (See appendix Table A 6). In particular, both trials suffered from threats to validity in regard to blinded outcome assessments for smoking cessation which was self-reported. Only DLCST used exhaled CO measurements at T0 and T1 for a more objective measurement to confirm quitting from smoking. Both trials, in addition, were subject to attrition bias due to loss to follow-up and non-participation in smoking surveys which was differential to treatment allocation and higher for participants in the control group.

10.9. Risk of bias for assessments of screening test characteristics of LDCT

Data on test characteristics of LDCT screening like the total number of performed LDCT screens, repetitive screens due to suspicious findings or false positive screening test with or without invasive procedures were in trials only provided from single arms. As these figures are not from randomized comparisons they are subject to bias typically encountered in observational cohort studies or case series. For these reasons no formal quality assessment of these outcomes was done.

11.Results of clinical effectiveness

11.1. Comparative effectiveness for critical and important outcomes

11.1.1. Lung cancer mortality

Seven RCTs (DANTE [43], DLCST [51], LUSI [56], MILD[58], NELSON [59,60], UKLS [70] and NLST [6,62] [70]) asessed the effects of LDCT screening compared with usual care and no screening or CXR screening (NLST [6,62]), and reported lung cancer mortality with ≥5 years of follow-up. [43],[51],[57],[70] As agreed with the purchaser of this HTA report during the scoping process, the NLST trial was included in this pooled analysis although the control group in this trial was screened with CXR. Borrowing external strength from indirect comparison the contribution of the NLST to the overall pooled effect of DLCT versus no screening is further explored in the network meta-analysis (see chapter 2.).

Figure 3 shows the pooled estimates of the seven RCTs comparing LDCT screening with no screening or CXR. The relative risk of death from lung cancer of LDCT compared with no screening or CXR was 0.80 (95%CI 0.72 to 0.88). The test of heterogeneity indicated no heterogeneity (I^2 = 0%), although single trial estimates of DANTE and DLCST were quite different from the remaining trials. Of note, the NELSON trial and NLST contributed roughly 75 % of weight to the pooled summary estimate as the number of patients included in these trials were higher than in the other studies.

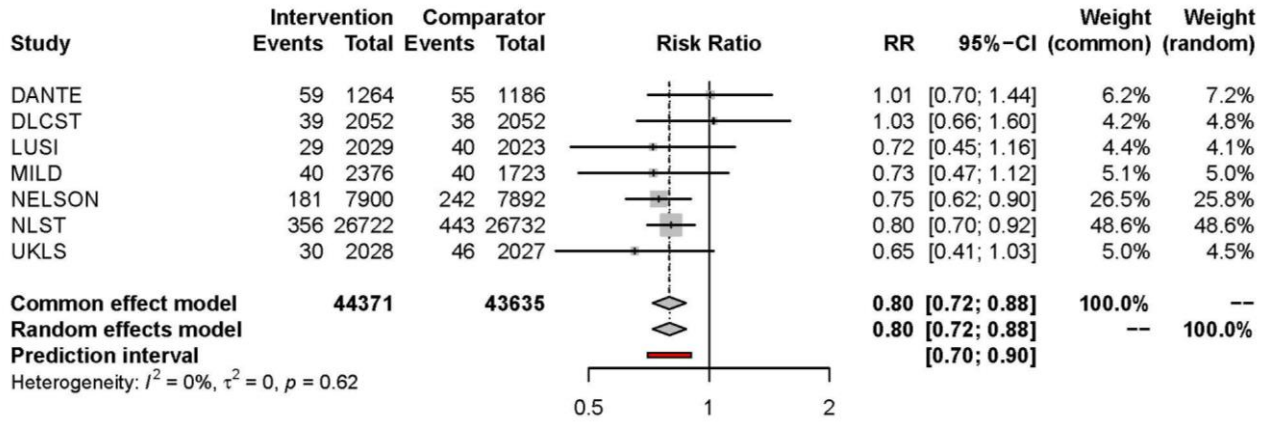


Figure 3 Forest plot of lung cancer mortality in trials with ≥5 years follow-up of LDCT versus no screening/CXR

11.1.2. All-cause mortality

Seven RCTs with ≥ 5 years of follow-up (DANTE [43], DLCST [51], LUSI [56], MILD [58], NELSON [59,60] UKLS [70]and NLST [6,62] [70] assessed the effects of LDCT screening compared with no screening or CXR screening (NLST [6,62]) on overall mortality (Figure 4). The NELSON trial provided overall mortality data only for men . The relative risk of death from all causes of LDCT compared with no screening or CXR was 0.96 (95% CI 0.92 to 1.00). The test of heterogeneity indicated no heterogeneity (I^2 = 0%). Again, the NELSON trial and NLST contributed roughly 75% of weight to the pooled summary estimate.

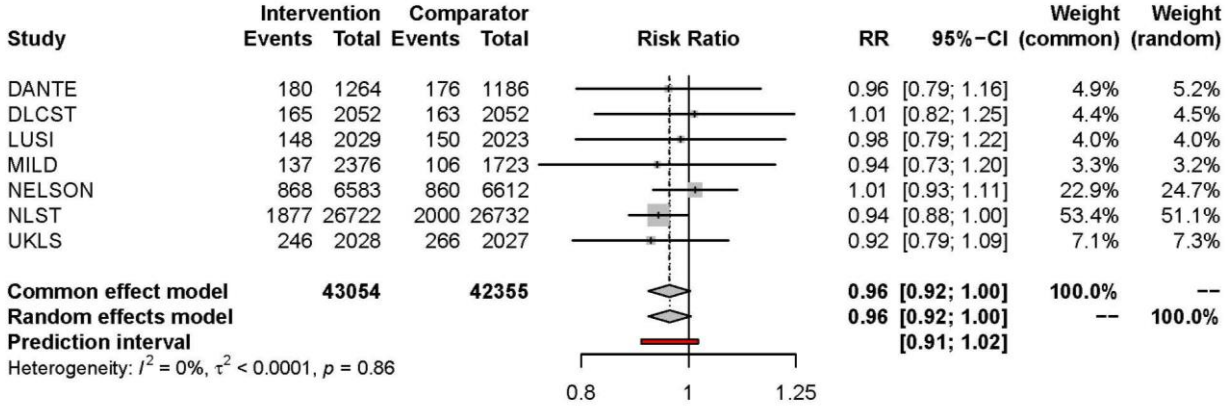


Figure 4 Forest plot of overall mortality in trials with ≥5 years follow-up of LDCT versus no screening/ CXR

11.1.3. Cancer detection

In total 7 trials with ≥ 5 years of follow-up provided data on cancer incidence and cancer stages at diagnosis and were pooled for a summary estimate. For reasons of consistency in relation to the definition of critical outcomes as delineated in the scope, we kept reporting of cancer stage limited to the trials with ≥ 5 years of follow-up. LDCT screening compared to controls (usual care/best available

care) was associated with an increased chance of lung cancer detection (pooled RR 1.35, 95% CI 1.14 to 1.60; I^2 = 67%; Figure 5).

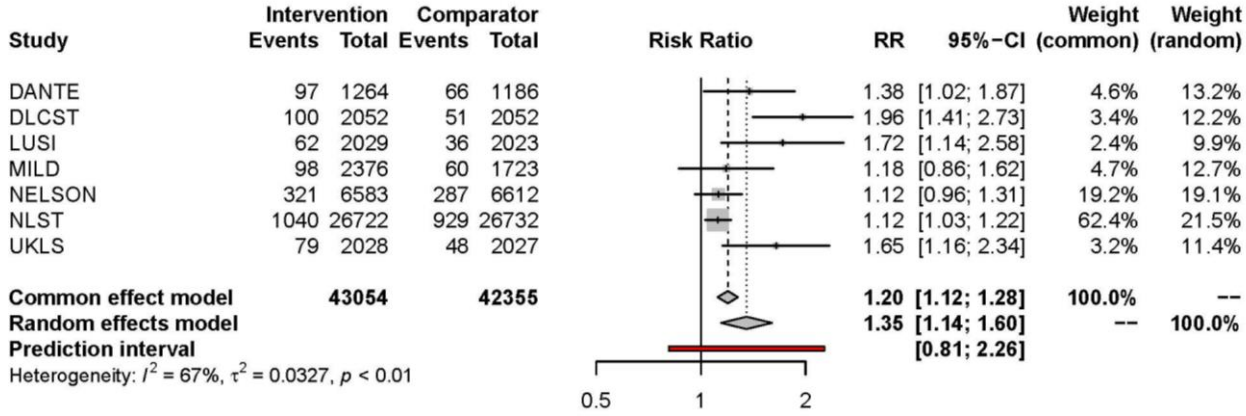


Figure 5 Forest plot for cancer diagnosis in LDCT versus no screening / CXR in trials with ≥ 5 years of follow-up

In a pre-specified additional analysis of all trials (i.e. including trials with outcome measure <5 years) comparing (Figure 6), the pooled relative risk for cancer detection of LDCT with no screening / CXR) was in the the same direction and basically little changed compared to the one based on trials with ≥5 years of follow-up (pooled RR 1.53, 95% CI 1.25 to 1.88; I^2 = 75%).

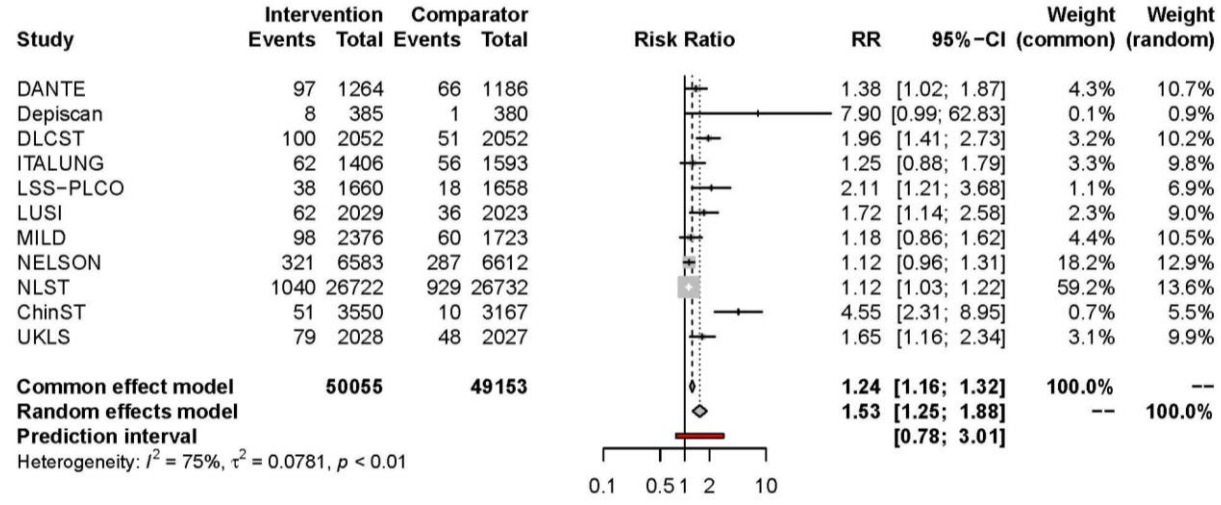


Figure 6 Forest plot for cancer diagnosis in LDCT screening versus no screening / CXR in all the trials

Moderate to high heterogeneity was observed in both meta-analyses (See Figures 5, and 6). As this additional analysis was not pre-specified we did not further explore heterogeneity.

11.1.4. Stage distribution

Seven trials with ≥ 5 years of follow-up provided relevant data for cancer stages as this was judged

sufficiently long to obtain mature data. For reasons of consistency in relation to the definition of critical outcomes as detailed in the scope, we kept reporting of cancer stage limited to trials with ≥ 5 years of follow-up. The forest plots in Figures 7 and 8 illustrate the change of lung cancer stage distribution between LDCT screening and control arms. Patients with LDCT compared to control were statistically significantly more likely to be diagnosed with lung cancers that were detected in earlier stages (I and II) (RR 2.69, 95% CI 1.94 to 3.74, I^2 = 80%) and less likely to be diagnosed at later stages (III and IV) (RR 0.79, 95% CI 0.72 to 0.86, I^2 = 0%). There was relevant heterogeneity for both summary estimates and for both summary findings single trial estimates from NLST were considerably different from the pooled summary estimates. When excluding NLST, heterogeneity was reduced for both estimates to some extent for diagnosis at stage I and II (RR 3.10, 95%CI 2.37 to 4.05; I^2 = 46%) and for diagnosis at stage III and IV (RR 0.83, 95%CI 0.70 to 0.97; I^2 = 38%).

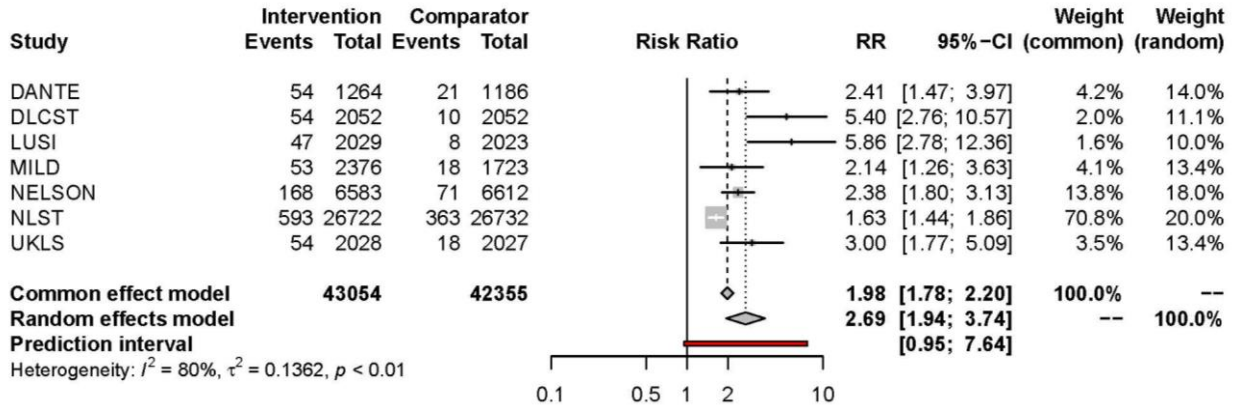


Figure 7 Forest plot for early stage (I or II) diagnosis in trials with ≥5 years follow-up of LDCT versus no screening/ CXR

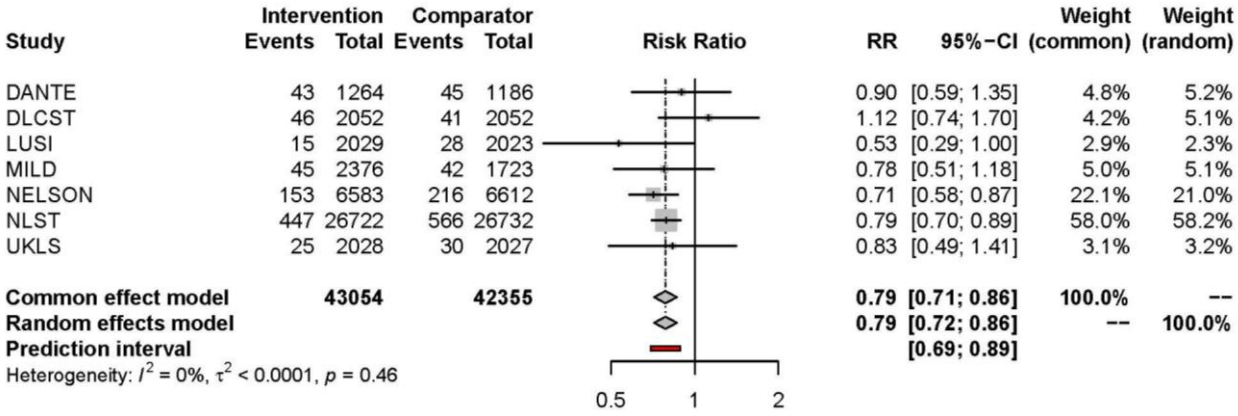


Figure 8 Forest plot of late stage (III or IV) diagnosis in trials with ≥5 years follow-up of LDCT versus no screening/CXR

In pooled analyses of all trials, irrespective of follow-up time, LDCT screening was associated with a statistically significant increase in early stage (I and II) cancer detection (RR 2.88, 95% CI 2.14 to 3.86, I^2 = 76%; Figure 9) and a statistically significant reduced likelihood of late stage (III or IV) cancer detection (RR 0.80, 95%CI 0.73 to 0.87, I^2 = 32%; Figure 10). There was considerable heterogeneity for

both summary estimates. Thus, summary estimates from all trials were not relevantly different from those of trials with ≥5 years of follow-up.

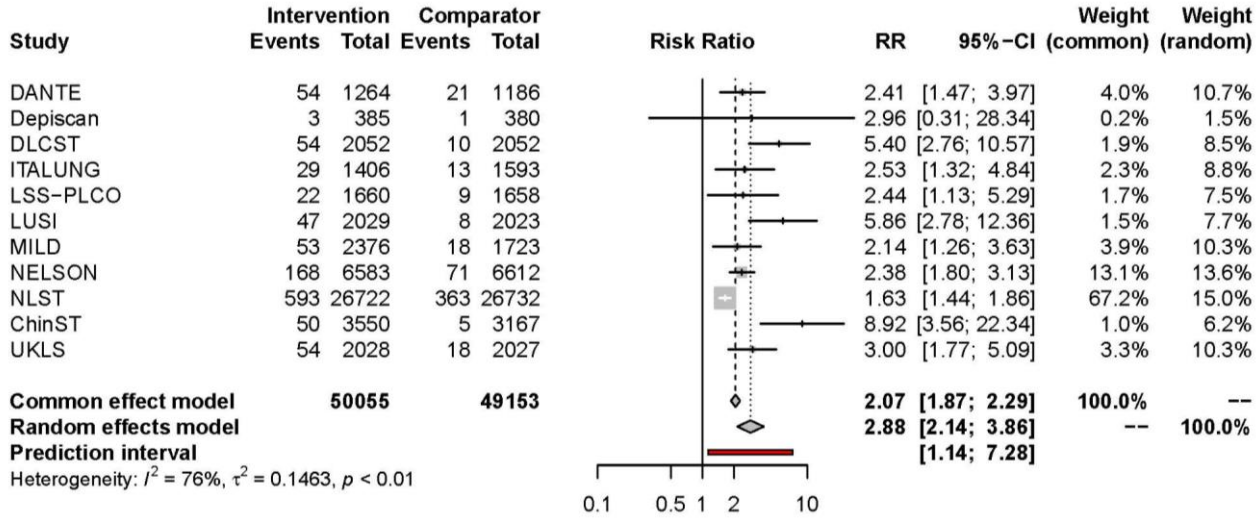


Figure 9 Forest plot of early stage (I or II) diagnosis in all trials of LDCT versus no screening/CXR

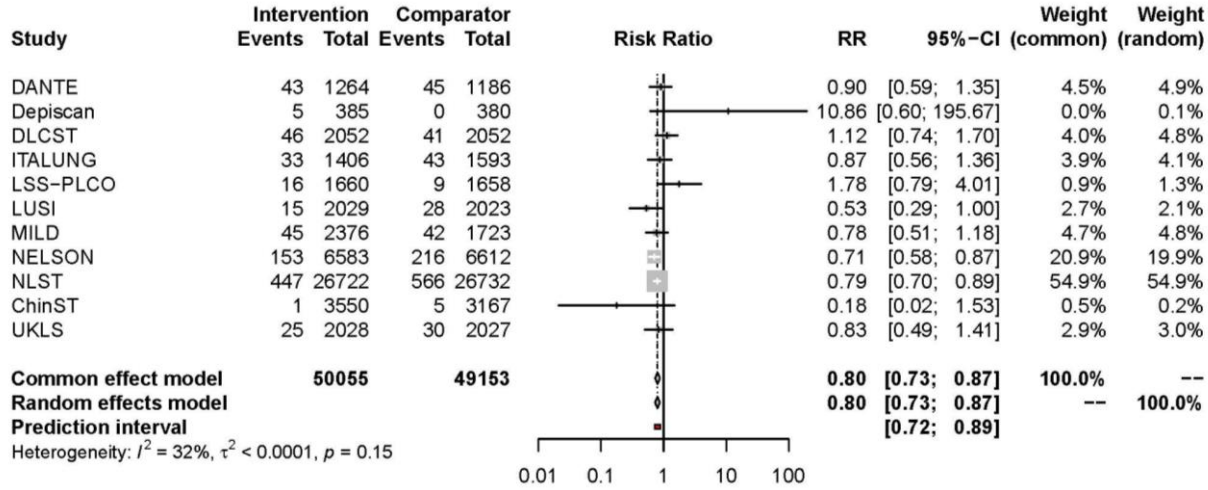


Figure 10 Forest plot of late stage (III or IV) diagnosis in all trials follow-up of LDCT versus no screening/CXR

11.1.5. Health-related quality of life and psychological consequences

Only three trials (DLCST, NELSON, and UKLS [63],[60],[51]) evaluated the psychological consequences of LDCT screening in trial participants in a representative subsample of trial participants (DLSCT n = 3,929; NELSON n = 1,466; UKLS n = 4,061) (Table 6). Only the NELSON trial [60] assessed the impact of LDCT screening on HRQoL of patients at short- and long-term follow-up. Three trials (DLCST, NELSON, UKLS[63],[60],[51]) assessed adverse psychological effects that may be associated with LDCT screening. All three trials had validity issues due to relative subjectivity of outcomes assessments, lack of blinding, and loss of follow-up. Trial subpopulations were generally compatible with the entire trial population in regard to demographic factors.

In DLCST [51], a lung cancer-specific questionnaire (COS-LC) was developed in 20 patients during

screening rounds and validated during subsequent screening rounds. COS-LC consists of nine psychosocial scales including four core scales ("Anxiety" (7 items), "Behaviour"(7 items), "Dejection" (6 items), and "Sleep" (4 items)) and five lung-cancer-screening-specific scales ("Self-blame" (5 items), "Focus on Airway Symptoms" (2 items),"Stigmatisation" (4 items), "Introvert" (4 items), and "Harm of Smoking" (2 items)). Only outcomes for the COS core scales are reported. In COS-LC higher scores indicate worse outcomes. In the DLCST [76], all randomized participants (n=4,104) in the LDCT screening and usual-care groups were invited annually to the screening clinic to complete the questionnaire consequences of screening lung cancer (COS-LC). During screening rounds completion rates for COS-LC for the LDCT group and usual-care group decreased substantially more in the control group and were 95.5% and 73.6%, respectively.

At the baseline screening round, mean scores for anxiety, dejection, negative impact on behavior and sleep were all lower in the LDCT group compared to the control group but differences in mean scores were small, likely not of clinical relevance and only statistically significant for 'dejection'. Prior to the second screening round mean score difference were more pronounced for all for scales in the LDCT compared to the control group and statistically significantly different for anxiety, negative impact on behavior and sense of dejection. During screening rounds 2 – 5 participants in the control group experienced statistically significantly more negative psychosocial consequences in seven of nine scales compared to the LDCT group.

The NELSON trial [83] examined HRQoL with the use of the Short Form questionnaire-12 items (SF-12) and the EQ-5D questionnaire in a random sample of 733 participants in each arm. Participants received questionnaires before randomisation (T0,) two months after baseline screening (T1) and at two years of follow-up (T2). Within EQ-5D the health status was rated on a visual analogue scale (range from 0 (worst) to 100 (best) health state). Anxiety was assessed with STAI-6 questionnaire, and the psychosocial stress (IES) was assessed with the cancer specific distress impact of event scale (range from 0-75; high scores indicating higher stress levels).

At baseline and at two years follow-up there were no statistically significant differences in mean scores for any scale between the LDCT and control group.

The UKLS [63]assessed the psychological consequences of LDCT screening and for health related quality of life HRQoL in 4,061 participants by use of the lung cancer distress scale (Cancer Worry Scale), the Hospital Anxiety and Depression Scale and the Hospital Anxiety and Depression Scale. Decision satisfaction was additionally assessed in this trial. Questionnaires were given to participants at baseline (T0), the 2-week post-scan result (T1) and at 2-year follow-up (T2). Follow-up results for cancer distress was only reported in low and high scorers and was therefore not abstracted. This was also the case for decisions satisfaction.

At two years follow-up, no statistically significant difference in the distress scores between the groups was found (LDCT 8.15 vs. control 8.10). Participants in the LDCT screening group had less anxiety compared with the control group at T1 (LDCT 3.67 vs. control 3.78), although this difference did not reach statistical significance. At the 2-year follow-up, participants in the LDCT screening group had statistically significantly less anxiety compared with the control group (LDCT 3.66 vs. control 4.02; $p \leq 0.001$). Furthermore, participants in the LDCT screening group had statistically significantly less depression than those in the control group at T1 (LDCT 2.53 vs. control 2.81; $p \leq 0.001$) and at T2 (LDCT 2.77 vs. control 3.01, $p \leq 0.01$).

At both T1 and T2, participants in the LDCT screening group had a statistically significantly higher satisfaction rate than the control group (T1: LDCT 42% vs. control 34%; T2: LDCT 40% vs. control 26%).

Study identifier	Participants Randomised (n) Participants in substudies [n] LDCT/control	Measures	Domain	LDCT; mean score (SD or 95% CI)	Control; mean score (SD or 95% CI)	P value for mean difference in score	Mean differences in scores between LCDT / control at study termination or last follow-up
DLCST [76,84]	4104 [2037/2042]	COS (high scores ≈ worse)	Prevalence round (Y1)				Screen positives (n = 179) not included in analysis
	[2052/1873]		Anxiety	1.48 (2.20)	1.61 (2.31)	<0.07	
			Neg. impact behavior	0.72 (1.78)	0.84 (2.08)	<0.05	
			Dejection	1.21 (1.99)	1.37 (2.17)	<0.03	
			Neg. impact sleep	0.63 (1.56)	0.70 (1.72)	=0.20	
			Incidence round (prior to second screening round, year 2)				
	[1884/1817]		Anxiety	1.50 (2.52)	1.71 (2.79)	<0.03	
			Neg. impact behaviour	1.76 (2.85)	2.02 (3.04)	<0.01	
			Dejection	1.61 (2.71)	1.88 (2.98)	<0.01	
			Neg. impact sleep	1.64 (2.47)	1.79 (2.57)	=0.10	
			Mean score increase from year 1 to year 5			(p values refer to score increase within each arm)	Mean difference Y1 to Y 5 Between LDCT and control
	[1825/1374]		Anxiety	–0.26 (–0.39 to –0.13)	0.25 (0.04 to 0.46)	LDCT: <0.0001 Control: <0.02	0.65 (0.41 to 0.89) <.0001
			Neg. impact behaviour	0.77 (0.63 to 0.91)	1.37 (1.13 to 1.60)	LDCT:<0.0001 Control:<0.0001	0.72 (0.45 to 1.00) <.0001
			Dejection	0.09 (0.04 to 0.22)	0.67 (0.44 to 0.90)	LDCT:=0.16 Control: <0.0001	0.73 (0.47 to 0.99) <.0001a
			Neg. impact sleep	0.83 (0.71 to 0.95)	1.53 (1.31 to 1.74)	LDCT:<0.0001 Control:<0.0001	0.78 (0.53 to 1.03) <.0001a

Table 6 Psychological consequences and Health related Quality of Life HRQoL (continued)

Study identifier	Participants (n)Randomised Participants in substudies [n] LDCT/control	Measures	Domain	Result LDCT; mean (SD or 95% CI)	Control; mean score (SD or 95% CI)	P value for mean difference in score	Mean differences in scores between LCDT / control at study termination or last follow-up
NELSON [83,77]	15,822	Baseline					
	[733/733]	EQ-5D (Higher scores ≈better)	QoL	79.19 (78.02 to 80.36)	78.50 (77.15 to 79.85)	n.s.	

		Anxiety Spielberger STAI-6 (high scores ≈ more anxiety)	Anxiety	33.27 (32.51 to 34.03)	33.75 (32.87 to 34.62)	n.s.	
		Lung-cancer specific distress (IES) (higher score ≈ worse distress)	Distress	4.05 (3.45 to 4.65)	4.02 (3.33 to 4.71)	n.s.	
	609 (89.3%)/322 (64.7%)	Year 2					
		EQ-5D	QoL	79.53 (78.35 to 80.71)	77.45 (75.95 to 78.95)	n.s.	
		STAI-6	Anxiety	32.67 (31.91 to 33.43)	33.42 (32.44 to 34.39)	n.s.	
		IES	Distress	3.72 (3.12 to 4.32)	4.03 (3.24 to 4.81)	n.s.	

Table 6 Psychological consequences and Health related Quality of Life HRQoL (continued)

Study identifier	Participants (n)Randomised Participants in substudies [n] LDCT/control	Measures	Domain	Result LDCT; mean (SD or 95% CI)	Control; mean score (SD or 95% CI)	P value for mean difference in score	Mean differences in scores between LCDT / control at study termination or last follow-up
UKLS [63,85]	4061 (2028/2027)	HADS anxiety and depression (high scores ≈ worse)	Baseline (T0)				
			HADS anxiety	3.72	3.67		
			HADS depression	2.66	2.61		
	[1553(82.3%) /1302 (65%.3)]		2 year (T2)				
			HADS anxiety	3.66 (3.52 to 3.80)	4.02 (3.86 to 4.19)	$p \leq 0.01$	

			HADS depression	2.77 (2.67 to 2.89)	3.01 (2.89 to 3.14)	p ≤ 0.01	

COS, consequences of screening; COS-LC, consequences of screening lung cancer; CWS, Cancer Worry Scale; HADS, Hospital Anxiety and Depression Scale; IES, impact of event scale; SD, standard deviation; SF-12, Short Form questionnaire-12 items; SF-36, Short Form questionnaire-36 items; STAI, State–Trait Anxiety Inventory; STAI-6, Spielberger State–Trait Anxiety Inventory, six-item Short Form; Y1, year 1; Y2, year 2; Y5, year 5. a Reported as change score with adjustment for multiple comparisons from year 1 to year 5

Table 6 Psychological consequences and Health related Quality of Life (HRQoL)

Overall, in two trials (DLCST and UKLS) assessing the psychological consequences of screening, participants in the LDCT group appeared to have somewhat lower mean scores for anxiety, depression and cancer related stress compared to the control group over the follow up time. Differences in scores were small and their clinical relevance is likely negligible .

11.1.6. Change in smoking behaviour

Lung cancer screening may increase participants' awareness of or fear from lung cancer disease and potentially be motivational to quit smoking. On the other hand screening may induce false reassurance not to quit or to delay the cessation attempts. For this important outcome we focused on long-term quitting rates, which had to be reported at baseline and at the end of the screening intervention in the entire smoking trial population in order to derive the most unbiased estimate of screening for lung cancer on smoking cessation rates. Selective inclusion of individuals irrespective of the screening procedures (e.g. LDCT or CXR in NLST) or focus on individuals with suspicions screening results will lead to biased smoking cessation estimates.

Four trials with ≥ 5 years of follow-up (DLCST [51], NELSON [59], [60], NLST [86] and UKLS [63]) provided data on smoking cessation during the screening phases of the trials. DLCST was the only trial that assessed smoking cessation behaviour in the entire trial population. The NELSON trial [87]selected two random samples of male smokers, who had received either only negative test results (n= 550) or one or more indeterminate test result (n=440), and who were sent a questionnaire on smoking cessation two years after randomisation. Female smokers were not included. As this sample was selected only among individuals in the LDCT arm and primarily based on screenees with suspicious findings the trial was excluded.

NLST [86] provided data on a subpopulation from the American College of Radiology Imaging Network (ACRIN) of 18,840 subjects, of which 8,358 individuals were current smokers at baseline. The focus in this report was on the motivational factors (e.g. positive LDCT or CXR findings) for smoking cessation. In addition, due to questionnaire modifications smoking, cessation data collection was stopped after 5.5 years of follow-up. No detailed data on smoking cessation by intervention groups was reported. In multivariate analysis no differences by study arm (LDCT vs. CXR) for smoking cessation of at least 6 months was found (HR =1.05, 95%CI 0.98, 1.14) after accounting for screening results and other covariates. For this reason the trial was excluded for the analysis of this important outcome.

Only DLCST [51] and UKLS [88] provided evidence on smoking cessation in all participants who were randomized to LDCT screening versus control (Table 7). The randomised evidence for smoking behaviour was rated as more prone to bias due to lack of blinding, loss to follow-up and self-reported

smoking data, with no additional measures as for examples CO single breath monitoring in one trial (UKLS).

DLCST [89], however, conducted CO measurements during baseline and the first year of screening. Participants in both arms received five annual visits including < 5-minute briefing for cessation advice in both trial arms. There was no statistically significant difference in ex-smoker rate at baseline between the screening group (25%) and the control group (no screening) (23%); p = 0.21. At the 1-year and 5 year follow-up, no statistically significant difference in quit rate between the screening group (11.3% and 42%) and the control group (10.4% and 40%) was observed (p = 0.47 and p = 0.075, respectively).

The findings from the DLCST [81] showed that screening with LDCT had no additional effect on participants' smoking status compared with the control group (no screening), however, smoking cessation markedly decreased in both arms in comparison between baseline and year 5 of follow-up.

In UKLS [88] participants who were smoking in the first questionnaire were eligible for analysis of smoking behavior during follow-up and received in both trial arms standard smoking cessation advice leaflets and a list of local National Health Service Stop Smoking services prior to randomisation. Smoking cessation was assessed by self-report at follow-up 1 and 2. Individuals who did not answer the questionnaire were included into the ITT analysis and were considered as smokers. At follow-up 1, completion rates were 527/758 individuals (70%) for the intervention arm and 479/786 individuals (61%) for the control arm. This decreased to 65% and 49% at follow-up 2, respectively. In the primary ITT multivariate analysis smoking cessation was more like in the LDCT compared to the control group 1.60 (1.17 to 2.18), but in a sensitivity analysis with further adjustment for cancer distress, recruitment site, gender, age, marital group, deprivation and experience of lung cancer the hazard ratio for smoking cessation of LDCT versus control was no longer statistically significant (1.16; 95%CI 0.65 to 1.33).

In conclusion, findings in the two trials that intended to evaluate smoking behavior change in relation to lung cancer screening at the broadest study population level, do not show that LDCT screening was associated with higher quit rates when compared to control. Results from remaining trials (NELSON and NLST) which examined selected populations are very likely subject to bias and were not considered for analysis in this report.

Study	LDCT Total randomsed smokers/ ex-smokers in brackets []	Missing	Control (no screening) smokers/ ex-smokers	Missing	LDCT vs. control	Time	Other results	Notes
DLCST [89]	2,052 [1545/507]	0	2,052 [1,579/473]	0	Smoker rates: 75% vs. 77% Ex-smoker rate: 25% vs. 23% (p = 0.21)	Baseline		Last observation carried forward used for missing data ITT analysis missing ≈smoking
	[1,051/806]	195	[937/713]	402	Smokers: 51% vs. 48% Ex-smoker rate: 43% vs. 43% (p = 0.909)	5 years	At 2-, 3-, 4-years similar results	
UKLS [63]	2,028 [758/75]	31	2,027 [786/36]	307	*Smoker rate: 50% vs.50% Ex-smoker rate; 68% vs. 32%	T1 two weeks after baseline scan results		ITT analysis missing ≈smoking
	[749/115]	261	[775/79]	398	*Smoker rate: 55% vs. 45% Ex-smoker rate: 59% vs.41%	2 years	NR	
ITT, intention to treat; NR, not reported in the identified study reports, * figures from sensitivity analysis accounting for additional confounders								

Table 7 Smoking status and smoking cessation

11.2. Non-comparative critical and important outcomes

11.2.1. False positives LDCT scans

False positives results from LDCT screening are important and key outcomes in screening programs in general, and particular for LDCT screening, because further scans and eventually invasive test procedures may be needed with the eventual negative consequences of complications, increased anxiety with other psychosocial consequences (like unnecessary sickness roles) and increased resource use. In the scoping process several key parameter for false positive LDCT were defined. It is very important to note, that the definition of a positive node or finding in LDCT varied between trials, and diagnostic work algorithms also differed (see chapter 4.1. on screening programme characteristics). In addition, the NELSON [59,60] and UKLS [63] trials made a distinction for the definition of indeterminate findings or interval imaging rate from false positives. In the NELSON [59,60] trial, the LDCT screening result was indeterminate if the volume of the largest solid nodule or the solid component of a partially solid nodule was 50–500 mm^3 or > 8 mm in diameter for non-solid nodule which was the case in 10.8% of all scans across three rounds of CT screening (2,629 out of 24,354 (e.i. 21,773 (89.4%) regulary scans plus 2581 (10.6%) follow-up scans needed to assess the volume doubling time of nodes of indeterminate size). In UKLS an interval scan was defined as protocol defined repeat scan after 3 months for a category 3 nodule. There were 472 (23.6%) subjects in this category. This rate constituted one of the two definitions of a false positive scan used in UKLS. The second definition of false positive scan in UKLS included all participants who were referred to the multi-disciplinary team or if an (additional) repeated scan was necessary within 12 months. There were 114 individuals in this category. In this HTA report indeterminate scans from NELSON and interval scans from UKLS are referred to as protocol defined indeterminate scans. Scans that did not constitute regular screening scans or protocol defined indeterminate scans are in this HTA report referred to as recall scans. Work up algorithms differed and were well defined in several trials but in the NLST participants with suspicious LDCT findings were referred to their private physicians and no specific work-up program was defined. For these reasons key characteristics of false positive tests are presented (as outline in the scope) separately for each trial. Some trials reported false positive scans and complications in relation to the number of scans, others in relation to the number of screened persons. For these reasons no formal pooling of individual trial results for positive scans or complications was done.

11.2.2. Number of false positive scans

A false positive scan depends on the definition of a positive node or findings in LDCT and work-up algorithms, which largely varied between trials (Table 3). In this report two definitions of a false positive scan were provided. The first definition is the rate of LDCT scans with *any* thorax related abnormality in relation to the total number of screening scans. The second definition, which can be considered more

relevant to participants in screening programs, is the rate expressed as the difference between the recall scans (that were outside of the regular screening schedule or protocol defined indeterminate scans) or work-up procedures and the screening detected lung cancer in relation to the total number of screened persons in the LDCT arms. The range of screening scans with any abnormality during the entire screening programs was wide: 4.5% in the MILD trial, 8.7% in LUSI trial, 10.2% in the NELSON trial, 24.2% in NLST and 47.7% in UKLS. The range of false positive scans according to the second definition was also large between trials and varied between less than 5% in DLCST (3.0%), MILD (3.6%), UKLS (4.0%) and the NELSON trial (1.2%) [59,60], and 37.4% in LUSI and 45.3% NLST trials [6,62] (Table 8). Trials with defined work-up algorithms had considerably lower false positive rates than the NLST trial[6,62] which had individuals with positive LDCT or CXR sent to their care providers.

When looking at the two largest trials (NELSON [59,60] and NLST [6,62]), which both had considerably different definitions of positive scans, the following facts are apparent: In the NELSON [59,60] trial (total of 24,354 scans) 59.4% (293 out of 493; 95% CI 54.8% to 63.9%) of the screen results from the work-ups were false positive for a follow-up period of 5.5 years with a positive predictive value of 40.6% for the total of 3 screening rounds. In total 24.5% (n=67) of individuals with a false positive results underwent an invasive diagnostic work-up.

In NLST [6,62], 94.9% of individuals in need of any form of recall scan or work-up in the LDCT group had a false positive scan over all screening rounds and were thus false positive. In NLST, no distinction was made between indeterminate findings, interval imaging findings or false positive findings. Rates of invasive work-up in NLST were 6.4% (1,706/26,722) for LDCT and 2.4% (636/26,732) for CXR. NLST also reported on complications in patients with false positive scans. Roughly 1.4% and 1.6% of participants in the LDCT and CXR group, respectively had at least one complication.

11.2.3. Number of false positives LDCT scans with invasive procedures

The number of false positive scans with invasive procedures ranged from 2.6% in MILD [57] to 9.6% in UKLS. [63] The rates of false positive scans with invasive procedures, however, should also be considered in relation to the rates or total number of false positive scans. These rates were low in DLST [51] , MILD [57], NELSON [59,60] and UKLS. [63]

11.2.4. Number of false-positive scans with complications

Four trials (DANTE, DLCST, NELSON and NLST[59,60], [43],[6,51,62]) reported major complications in individuals who underwent any surgical or other invasive diagnostic work-up procedure independent of whether the work-up resulted in a cancer diagnosis or not, and two of these trials (DANTE and NLST [43],[6,62]) provided data for both the LDCT screening and the control arm.

In the DANTE trial[43] 45 patients in total underwent a major surgical procedure for suspected lung cancer,

and the diagnosis was confirmed in 39. Six of 45 (13.3%) individuals in the LDCT group and 3 of 20 (15%) in the control group had major surgical procedures for benign disease. In the DLCST trial [51] 25 participants in the LDCT arm had an invasive work-up with a total of 40 invasive procedures. Of those, 10 invasive procedures in 8 patients resulted in benign disease. In NELSON (according to reference [60]) there were 458 out of 7,582 individuals in the LDCT arm who had at least one positive scan and of those 200 were diagnosed with lung cancer. A total of 273 participants had one or more false positive scans, and 61 of 273 participants (24.5%) had an invasive work-up that resulted in a benign disease. In NLST [6] there were 457 invasive procedure in the LDCT arm which turned out to be for benign lesions and of those 61 complications were noted. There were 107 invasive procedures in the CXR arm which turned out to be for benign lesions and of those 16 complications were noted.

In trials reporting data from both groups e.g. the DANTE trial [43] and the NLST [6,62], rates of invasive procedures were higher in the LDCT screening arm than in the control group; in the DANTE trial[43], 28.6% in the LDCT group and 19.3% in the usual-care group, and in the NLST trial12.0% in the LDCT group and 9.0% in the CXR group.[43],[6, 62]

Data on complications from false positive LDCT was, however, very scarce. Three trials (DLCST [51], NELSON [59,60], NLST [62]) reported data on the complications following invasive diagnostic work-up of false positive LDCT or CXR screening. In DLST, there were 2 and 0 complications during invasive work-ups of false positive scans, in the LSS-PLCO there were 2 and 1 complications. In NLST [6], there were 457 invasive procedure in the LDCT arm for benign lesions and of those 61 complications were noted. There were 107 invasive procedures in the CXR arm for benign lesions and of those 16 complications were noted.

11.2.5. Number of indeterminate scans

Not all trials reported on indeterminate scans. The NELSON [59,60] and UKLS [63] trials had explicit definitions of indeterminate findings in their work-up algorithms. As a consequence in the NELSON trial [59,60], the LDCT screening result was indeterminate in 10.8% (2,629 out of 24,354) of all scans across three rounds of LDCT screening. As outlined in table 3 in the NELSON trial[59,60] an indetermined scan was defined as a volume of the largest solid nodule or the solid component of a partially solid nodule of 50–500 mm^3 or > 8 mm in diameter for non-solid nodule. Indeterminate results led to repeat scans after 6–8 weeks or after 12 months, depending on the nodule size and screening round in order to determine the final result as positive or negative.

In UKLS [63], an indeterminate scan was defined as a positive scan that was referred to the multidisciplinary team (MDT) and/or repeated imaging scans before 12 months and did not results in lung cancer diagnosis. The UKLS reported that the interval imaging rate for the category 3 (larger, potentially malignant) nodules was 23.2%.

11.2.6. Number of follow-up investigations (invasive and non-invasive) following LDCT screening

The rates of additional scan investigations and the rates of noninvasive additional CT examinations outside from the established screening time table also varied greatly between trials with a range from 2.1% in the NELSON [59,60] trial to 47.7% in the UKLS trial[63]. Rates of invasive procedures per screened individual were low in DLCST (0.5%), MILD (0.6%), NELSON [59,60] (1.0%) and UKLS [63] (0.6%) but increased to 6.4% in the NLST [6,62] and 11.4% in the DANTE [43] trials.

11.2.7. Number and type of lung cancer treatment

Information on the type of cancer treatment was generally scarce in trials. DANTE [43] reported that 77 of 1,276 individuals in the LDCT and 31 of 1196 individuals in controls underwent surgery. DLCST provided the most detailed information of all trials. Of 68 and 24 lung cancers detected during screening in the LDCT and control groups, 51 in LDCT and 8 in controls were treated with surgery, and 17 and 16 lung cancers were treated with combined radiotherapy and chemotherapy, respectively. In ITALUNG [53] of 21 lung cancers, 13 were treated with surgery, 4 with chemotherapy and 13 with radiotherapy. In MILD [57] 41 of 49 lung cancers in the LDCT group were amenable to surgery. In UKLS [63], 35 lung cancers in the LDCT group were treated with surgery, 11 with chemotherapy and 5 with radiotherapy.

11.2.8. False negatives

False negative screening results are also of concern as this may lead to false reassurance in the absence of a pathology, which in reality exists and as a consequence can lead to delayed diagnosis and poor outcome. Four trials (DANTE,[71]MILD,[90]NLST [91] and NELSON [92]) indicated a sensitivity of CT screening from 69% to 94%. Three trials DANTE [71] MILD [90] and NELSON [92] gave rates of false negative scans (defined as scans that remained negative after 12 months follow-up) ranging from 0.1% in NELSON [91], 0.002% in MILD [90] bi-annual screening, 0.003% in MILD [90] annual screening to 1.3% in the DANTE trial [71]. For NELSON [59], the negative predictive value over the three screening rounds was 99.8% (95%CI 99.8% –99.9%).

12.Overdiagnosis

Overdiagnosis relates to the fact that in cancer screening programmes in particular early cancers are detected and treat which in the absence of screening would have never become apparent and treated. Overdiagnosis is particularly critical if diagnosis and treatment of cancer are done in individuals who are likely not to benefit from screening due to limited life expectancy in particular from comorbidities. There is a simple approximative approach to overdiagnosis estimation, one which emphasizes the patients' perspective (denoted P_s) and one that emphasizes the public health perspective (denoted P_a). [93] The excess incidence P*s* is calculated as the difference in cumulative incidence of lung cancer

between the LDCT and control arms, and express it as a ratio relative to the cumulative incidence of screen-detected lung cancers. P_s reflects a measure for the probability that a participant's LDCT screen–detected cancer would not have become clinically apparent during the active screening program if LDCT screening had not been performed. Based on this method, estimated overdiagnosis rates were 25.4% (−11.3% to 64.3%) in the LUSI trial [94] and 18.5% (95% CI 5.4% to 30.6%) in NLST.[93] Over a median follow-up time of 11.3 years for cancer incidence the overdiagnosis rate in NLST declined to 3.1%.[95] P_a has the same nominator, the denominator, however, is the total number of lung cancers diagnosed in the LDCT arm. P_a reflects the fraction of all lung cancers diagnosed during the active screening phase that would not have been diagnosed in the abscence of LDCT screening. For both methods 95%CI are calculated via bootstrapping. In NLST the fraction of lung cancers diagnosed during the active screening phase P_a was 11.0% (3.2% to 18.2%). [93] In the NELSON trial the fraction *Pa* at 4.5 years of follow-up and conclusion of the active screening phase was estimated at 19.7% (95% CI, −5.2% to 41.6%) 5.5. years after the closure of the screening program. At 11 years of follow-up the overdiagnosis rate, however, was reduced to 8.9% (95%CI -18.2 to 32.4%). These figures [61] indicate that the presence of lead time bias is substantial and requires extended follow-up periods for more appropriate assessments of overdiagnosis.

In LUSI further modeling was done and [94] the maximum likelihood estimates of mean preclinical sogiorn time (MPST) for all lung cancer types was additionally estimated: 47.5% (43.2%, 50.7%) of screen-detected tumors had a lead time ≥4 years, 32.8% (28.4%, 36.1%) a lead time ≥6 years and 22.6% (18.6%, 25.7%) a lead time of ≥8 years. Thus, about 43%, 33%, 23%, 16% and 11% of screen-detected tumors would in the absence of screening have remained in a preclinical phase over, 4, 6, 8, 10 and 12 further years, respectively.

Study	LDCT arm n	Total number of planned screening scans (% adherence)	Number of screening scans with any abnormal finding∞ or protocol defined indeterminate scans (n) per total scans (%)	Number of recall scans£ or work-ups (n) per total scans Ц(%)	Screening detected lung cancers (n)	False positive scans	False positive rate #	Number (n) and rate of invasive procedures per screened individual	Number (n) and rate (%) of individuals with unconfirmed cancer and invasive procedures per recall scan or work-up
DANTE [43], [71]	1,264	3,612 (93.7%)	562 (15.6%)	355 (9.8%)	66	289	22.8%	144 11.4%	17 4.8%
DLCST [96], [97]	2,052	9,800 (95.5%)	1,029 (10.5%)	302 (3.1%)	241	302	3.0%	10* (0.5%)	10* (6.2%)
LUSI [55] , [56]	2,029	9,405 95.1%	816 (8.7%)	816 (8.7%)	58	758	37.4%	158 (7.8%)	23 (2.8%)
MILD (annual &biannual)± [57]	2,303	10,038 88.7%	524 (4.5%)	150 (1.3%)	65	88	3.6%	13 (0.6%)	4 (2.6%)
NELSON [61] [60]	7582	24,353≠ (85.8%)	2629 (10.8%)	493 (2.0%)	200	293	1.2%	67 (1.0%)	NR¥
NLST LDCT:	26,722	75,126 (95.0%)	18,146 (24.2%)	12,757 (17.0%)	649	12,108	45.3%	1,706 (6.4%)	618 (4.8%)
CXR:[6]	26,732	73,470 (93.0%)	5,043 (6.9%)	4,211 (5.7%)	279	3,922	14.7%	636 (2.4%)	264 (5.5%)

UKLS [63]	2,028	1,994 (95.9%)	951 (47.7%)	114 (5.7%)	42	72	3.6%	12 (0.6%)	11 (9.6%)

[∞] Refers only to any abnormall findings in LDCT of the thorax

≠ In the Nelson trial[60] the algorithm for the definition of a positive node requested the conduct of follow-up scan to determine the volume doubling time of indeterminately sized nodules . These scans were subsummized under the total number of planned screening scans.

[£] Scans that did not constitute regular screening scans or protocol defined indeterminate scans are in the HTA report referred as recall scans

Ц The number of recall scans or work-ups as defined in individual trials are summarize in this columm.

Interval scans

[#] Definition of false positive rate [recall scans/work-ups – screening detected cancers] / total number of individuals

*Only available for baseline screening,

± Only reported for the first 7 of nine planed rounds (T0 – T6) for annual and the four rounds of biannual screening (T0-T3). The median follow-up in MILD for these figures are 6 years ± 1.7 follow-up for annual scans and 3 years follow-up± 1.7 for biannual scans over a median follow-up of 7.3 years.

[¥]not reported.

DANTE reported 17 and 5 complications from work up in false positive LDCT and control groups, DLST 2 and 0, LSS-PLCO 2 and 1 and NLST reported 44 and 8 complications from work up in false positive LDCT and CXR, respectively. NR not reported

Table 8 Number of false positive LDCT scans, rates of additional scans, rates of invasive work-up procedures and rates of invasive procedures in false positive scans

13.Network meta-analysis

In the primary analysis on the effectiveness of LDCT screening for lung-cancer, trials were pooled irrespective whether the control groups received no screening or whether a CXR was part of the control strategy as in the NLST. In a sub-analysis that was pre-specified in the scoping document and in agreement with the contracting body, the relative effectiveness of LDCT and CXR screening was further elucidated by indirect comparisons of LDCT, CXR and no screening in a triangular network meta-analysis that was restricted to the critical outcomes lung cancer and overall mortality.

The network meta-analysis is based on the trials that were identified in the primary analysis. In our updated literature search no additional trials that compared CXR versus no screening were found in addition to the ones identified in the HTA report by Snowsill et al. [1]

13.1. Characteristics of included studies

There were 10 trials eligible for the network meta-analysis, 7 trials from the primary analysis where of those 6 trials (DANTE [43]DLCST [51]LUSI [56] MILD [58] NELSON[59,60] and UKLS [70] compared LDCT to usual care with no screening, and one trial (NLST) [6] compared LDCT to CXR. Three trials the Czech trial[98], Mayo clinics trial [99], and the PLCO trial[100] compared CXR to usual care with no screening. Baseline characteristics of the seven trials comparing LDCT to no screening and CXR, respectively, are provided in tables 2 and A2 and have been described above.

Study identifier	Country	Recruitment time	Screening programme	Comparator	Sample size	Eligible age range per protocol (years)	Number of screening rounds	Screening times and interval (years)	Duration of follow-up
Kubík and Haerting [98]	Czech Republic	June 1976 to June 1977	CXR	Usual care	n = 6,364	40–64	Frequently = 9 vs. less frequently = 5	Frequently: every 6 months for 3 years then once in years 4, 5 and 6 vs. less frequently: prior to randomisation, at 3, 4, 5 and 6 years	15 years
Mayo [99]	USA	August 1971 to NR (screening ended July 1976)	CXR, sputum cytology	Usual care (recommendation for an annual CXR and sputum cytology)	n = 11,001 (planned 10,000)a	> 45	18	4 months	6 years
PLCO (for sensitivity analysis only)[100]	USA	1993 to 2001	CXR	No screening	n = 154,901 (NLST eligible subgroup, n = 30,321)	55–74	4	Annually	6 years (for NLST eligible subgroup)

NR, not reported; PLCO, Prostate, Lung, Colorectal and Ovarian cancer screening trial.
a Fontana et al. and Marcus et al [99]. report; n = 10,933.

Table 9 Baseline characteristics of randomised controlled trials of CXR screening versus no screening included into the network meta-analysis for lung cancer mortality outcome

In Table 9 baseline characteristics of the additional three trials comparing CXR to usual care with no screening are given, which were considered for the network meta-analysis for lung cancer mortality outcomes. These trials were conducted in the Czech republic and in the US [98-100]. The Czech trial [98] recruited participants from a general health prevention examination, the Mayo Clinic Trial [99] from the institution's outpatient clinic and the PLCO trial [100] from 10 centers in the US. All trials included current or heavy predominantly male, ex-smokers of similar age and thus represent high risk populations. The number of randomized participants ranged from 6,246 to 10,933. The PLCO [100] trial recruited 154,901 individuals who were smokers and non-smokers and reported in an ancillary analysis on lung cancer deaths in the trial sub-population of heavy smokers that fulfilled recruitment criteria for the NLST trial. For the present analysis we included the subpopulation (n= 30,321) from PLCO [100] that was followed over a period of 6 years. A post-hoc analysis concluded that the trial was heavily underpowered to show a statistically significant effect for this subpopulation. In a sensitivity analyses we therefore excluded this trial from the network meta-analysis.

Screening intervals were markedly different between trials. In the Czech trial individuals received CXR screening at baseline and every 6 months from years 1 to 3, and then every year from 3 to 6 years onwards. In the Mayo Clinic Trial CXR was conducted every 4 months for a period of 6 years. In the PLCO trials participants received a CXR at baseline and then every year for a period of 3 years. CXR readings also differed between trials. The Czech trial and the Mayo Clinic Trials used double readings (with referral of individuals with X-ray abnormalities at the reader's discretion in the Czech trial, the Mayo Clinic Trial did not specify a work-up algorithm) and in the PLCO trial participants with a pathology finding in X-rays were referred to patients' primary care providers for further work-up. Details of baseline characteristics of trial participants, recruitment, screening procedures, judgements of CXR and adherence to screening programmes are provided in the appendix in Tables A 8 to A 10.

Risk of bias for the three trials comparing CXR to no screening is presented in Table A 11. Generally quality of these trials was judged to be lower compared to trials of LDCT screening. The Czech trial [98] and Mayo trial [99] were both judged at high risk of bias due to lack of specification of randomisation, high risk of contamination, lack of documention of blinding for outcome assessment and high risk of attrition bias.

13.2. Pooled estimates from network meta-analysis

The network meta-analysis comprises 10 trials with a follow-up ≥ 5 years, 6 trials comparing LDCT to no screening, one trial comparing LDCT to CXR, and 3 trials comparing CXR to no screening. Figure 11 shows the triangular network formed by the respective trials and comparisons of screening interventions.

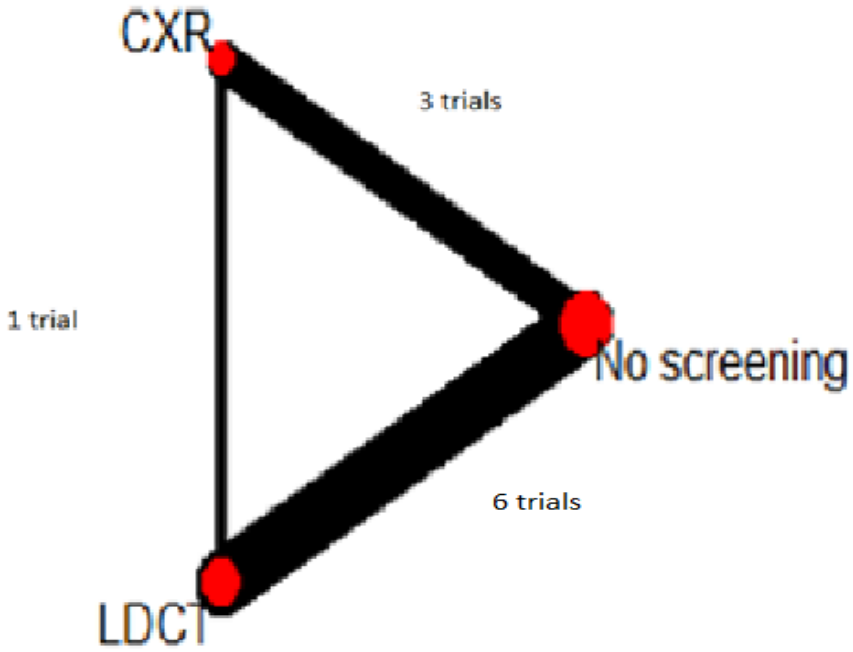


Figure 11 Triangular network of randomised controlled trials reporting lung cancer mortality and overall mortality data of LDCT with no screening or LDCT screening with CXR screening, and trials comparing CXR with no screening (for overall mortality only 2 trials comparing CRX with no screening provided data)

In the network analysis, the league table of pooled risk ratios using a random-effects model for the direct and indirect comparisons of trials comparing LDCT with CXR, trials comparing LDCT with no screening, and trials comparing CXR with no screening are provided for the critical outcome of lung cancer mortality and overall mortality. In Table 10 the RRs to the left of the reference comparison represent the pooled results of the direct and indirect comparisons and estimates, RRs to the right represent the results for the direct comparisons. Thus, pooled evidence (indirect and direct comparisons) indicate that CXR screening compared to LDCT is associated with an increased risk of death from lung cancer (RR 1.28, 95% CI 1.15 to 1.43] an estimate not much different from the direct estimate (RR 1.24 95%CI 1.08 to 1.43]. Likewise, pooled overall evidence indicates that LDCT compared to no screening is associated with reduced risk of lung cancer (RR 0.81 95%CI 0.73 to 0.90), an estimate that is not much different from the direct comparison (RR 0.79 95%CI 0.73 to 0.90). Q statistics for the assessment of homogeneity and consistency indicated a low probability of inhomogeneity or inconsistency with an overall p-value of 0.41, and within and between design p values of 0.33 and 0.63, respectively. Estimates of direct and indirect comparison for overall mortality indicate no statistically significant benefit of any screening measure. For this estimate data was only available of two trials comparing CRX with no screening.

Lung cancer mortality		
CXR	1.24 [1.08, 1.43]	1.05 [0.96, 1.16]
1.28 [1.15, 1.43]	**LDCT**	0.79 [0.69, 0.90]
1.04 [0.96, 1.13]	0.81 [0.73, 0.90]	**No screening**
Overall mortality		
CXR	1.07 [1.00, 1.13]	1.01 [0.94,1.09]
1.05 [1.99, 1.11]	**LDCT**	0.98 [0.92, 1.05]
1.03 [0.97, 1.09]	0.97 [0.92, 1.03]	**No screening**
Estimates on the lower left of reference comparisons represent pooled results of direct & indirect evidence (NMA), those on the upper right direct comparisons.		

Table 10 League table with risk ratios of death from lung cancer and overall mortalty comparisons of LDCT with CXR or with no screening

Due to the very small number of nodes and trials in the network, no ranking of the three screening strategies was done, as a ranking in such a situation is likely to provide inconsistent results.

13.3. Sensitivity analysis in network meta-analysis

We performed one sensitivity analysis by excluding the PLCO trial in which the RR estimates were similar with no change in summary estimates of direct and indirect comparison (Table 11).

Lung cancer mortality		
CXR	1.24 [1.08, 1.43]	1.12 [1.00 - 1.26]
1.31 [1.17, 1.46]	**LDCT**	0.79 [0.69, 0.90]
1.08 [0.98, 1.20]	0.83 [0.74, 0.92]	**No screening**
Estimates on the lower left of reference comparisons represent pooled results of direct & indirect evidence (NMA), those on the upper right direct comparisons.		

Table 11 League table with risk ratios of death from lung cancer in direct and indirect comparison of LDCT with CXR or with no screening with the exclusion of the PLCO trial

13.4. Risk of bias in network meta-analysis

The risk of bias in the network meta-analysis was assessed with the CineMa tool. As outlined in Figure 12 risk of bias was low in all domains with the exception of within study bias and precision for the comparisons of CXR versus no screening.

Comparison	Number of studies	Within-study bias	Reporting bias	Indirectness	Imprecision	Heterogeneity	Incoherence	Confidence rating	Reason(s) for downgrading
CXR:LDCT	1	No concerns	Low risk	No concerns	No concerns	No concerns	No concerns	High	
CXR:No screening	3	Some concerns	Low risk	No concerns	Major concerns	No concerns	No concerns	High	*
LDCT:No screening	6	No concerns	Low risk	No concerns	No concerns	No concerns	No concerns	High	

* ["Within-study bias","Imprecision"]

Figure 12 Risk of bias assessment of randomised controlled trials reporting lung cancer and overall mortality data of LDCT with no screening, or LDCT screening with CXR screening, and trials comparing CXR with no screening

13.5. Summary of findings

In total 13 RCTs were identified of which 11 trials compared LDCT with no screening, and three trials compared LDT with CXR. Of these trials, 7 had a follow-up of ≥5 years and were considered relevant according to the scope of this HTA report for further evaluations and data synthesis. These trials had a total of 88'006 participants included, where 2 trials contributed roughly 75% of weight in pooled analyses of the critical outcome of lung cancer deaths. Characteristics of included populations varied between trials, as definitions of high risk populations for lung cancer were not uniform. All trials with the exception of four trials had beside the baseline screening, at least 3 following screening rounds, whereas the latter trials only had a baseline or one additional screening round. Definitions of a positive scan varied widely across studies in terms of nodule sizes or volume, and in work-up algorithms for non-normal CT scans. All but two trials used diameter size of non-calcified nodules with or without morphological suspicious malignancy aspect as a criteria for a positive finding. Critical diameter size again varied substantially between trials. Only two trials used volume parameters and volume doubling time as a criteria for a positive screening finding. These discrepancies had major impact on the rates of the false positive LDCT findings. Reporting of outcomes that were considered critical or important, during the scoping process of this HTA report, varied greatly between trials. The majority of trials did not report for example on the psychological outcomes like quality of life, stress or anxiety.

Three ongoing trials with roughly an additional 36,000 participants for LDCT screening versus no screening were identified. Completetion of these trials remains unclear but may be expected between 2022 and 2023.

13.6. Critical outcomes

LDCT screening compared with no screening or CXR was associated with a reduced **risk of death from lung cancer** (RR 0.80 , 95%CI 0.72 to 0.88; I^2 = 0%) but with **death from all causes** not being statistically significantly reduced (RR 0.96, 95%CI 0.92 to 1.00; I^2 = 0%). The NELSON trial and NLST contributed roughly 75% of weight to the pooled summary estimate for both critical endpoints. The I^2 test indicated no heterogeneity for both outcomes but it is known that its power is low. Visusal inspection of forest plots for both summary findings do not appear to be without heterogeneity. Risk of bias for these critical outcomes was considered to be low. Table 12 provides a summary of the expected effects of LDCT screening versus no screening or CXR for individuals with at least a 5 year of follow-up using the GRADE approach.

In a triangular network meta-analysis the relative effectiveness of LDCT and CXR screening was further elucidated by indirect comparisons of LDCT, CXR and no screening that was restricted to the primary endpoint of lung cancer mortality. Five trials (DANTE [43]DLCST [51] LUSI [55] MILD [57], NELSON, and UKLS [70])

compared LDCT to usual care with no screening, and one trial (NLST [6,62]) compared LDCT to CXR. Three trials the Czech trial [98], Mayo clinics trial [99], and the PLCO trial[100] compared CXR to usual care with no screening. Pooled evidence (indirect and direct comparisons) indicated that CXR screening compared to LDCT is associated with an increased risk of death from lung cancer (RR 1.28, 95% CI 1.15 to 1.43] - an estimate not much different from the direct estimate (RR 1.24 95%CI 1.08 to 1.43). LDCT compared to no screening was associated with a reduced risk of lung cancer in the pooled estimate of the direct and indirect comparison (RR 0.81 95%CI 0.73 to 0.90), an estimate not much different from the direct comparison (RR 0.79 95%CI 0.73 to 0.90). There was a low probability of inhomogeneity or inconsistency in the network and risk of bias according to the CineMA approach was considered to be low. No statistically significant difference from any direct and indirect comparison was found in the network meta-analysis for overall mortality.

13.7. Important outcomes

Patients with LDCT compared to control were statistically significantly more likely to be diagnosed with lung **cancers that were detected in earlier stages** (I and II) (RR 2.69, 95% CI 1.94 to 3.74, I^2 = 80%) and less likely to be diagnosed at later stages (III and IV) (RR 0.79, 95% CI 0.72 to 0.86, I^2 = 0%). Risk of bias was considered to be low for this important outcome.

Only three trials (DLCST, NELSON, and UKLS [63],[60],[51]) evaluated the **psychological consequences of LDCT screening** in trial participants, two in a representative subsample of trial participants. Only the DLSCT trial evaluated the entire population. All three trials had validity issues due to relative subjectivity of outcomes assessments, lack of blinding, and loss to follow-up. No uniform picture in terms of psychological consequences from screening with LDCT can be drawn. In DLCST following the first and prior to the second screening round, mean scores for anxiety were lower in the screening group, but likely not clinically relevant. During screening rounds 2 – 5 participants in the control group experienced statistically significantly more negative psychosocial consequences in seven of nine health scales compared to the LDCT group. The NELSON trial [83] examined HRQoL with the use of the Short Form questionnaire-12 items (SF-12) and the EQ-5D questionnaire in a random sample of 733 participants in each arm. Anxiety was assessed with STAI-6 questionnaire and psychosocial stress (IES) was assessed with the cancer specific distress impact of event scale. Participants received questionnaires before and two months after baseline screening and at two years of follow-up. At baseline and at two years follow-up there were no statistically significant differences in mean scores for any scale between the LDCT and control group. The UKLS [63] assessed the psychological consequences of LDCT screening and for health related quality of life HRQoL in 4061 participants by use of the lung cancer distress scale (Cancer Worry Scale), the Hospital Anxiety and Depression Scale and the Hospital Anxiety and Depression Scale at baseline, 2-week post-scan result

and at 2-year follow-up and no statistically significant difference in the distress scores between the two groups were found. Participants in the LDCT screening group had less anxiety and depression compared to the control group following the first screening and at 2 years follow-up.

Only two trial (DLCST and UKLS) investigated **smoking behavior** in the entire trial population and found no difference in smoking cessation between the LDCT and control groups.

False positive scans depend on the definition of a positive node or findings in LDCT and work-up algorithms which largely varied between trials. The range of any found thorax anormality during screening programs was wide and between 4.5% in the MILD trial and 47.5% in the UKLS trial.

The range of false positive scans (defined as [recall scans or work-ups – screening detected lung cancers] / screened individuals) was also large between trials and varied between1.2% in NELSON and 45.3% in the NLST trial [6,62]. Trials with defined work-up algorithms had considerably lower false positive rates than the NLST trial [6,62] which had individuals with positive LDCT or CXR sent to their care providers. The two largest trials (NELSON [59,60] and NLST [6,62]) had considerably different definition of positive scans. Whereas NELSON had an additional definition of an indeterminate scan, NLST had not. In the NELSON trial [59,60] (total of 24,354 scans) 59.4% (293 out of 493; 95% CI 54.8% to 63.9%) of positive screening results were false positive for a follow-up period of 5.5 years. In NLST [6,62] 94.9% of individuals in need of any form of recall scan or work-up in the LDCT group had a false positive scan over all screening rounds and were thus false positive. In NLST no distinction was made between indeterminate findings or interval imaging findings and false positive findings.

The number of **false positive scans with invasive procedures** ranged from 2.6% in MILD [57] to 9.6% in UKLS [63].

Four trials reported major **complications in individuals who underwent <u>any</u> surgical or other invasive diagnostic work-up procedures** independent of whether the work-up resulted in a cancer diagnosis or not, and two of these trials (DANTE and NLST)[43],[6,62]provided data for both the LDCT screening and the control arm. NLST also reported on complications in patients with false positive scans. Roughly 1.4% and 1.6% of participants in the LDCT and CXR group, respectively had at least one complication. Data on **complications from false positive LDCT** was, however, very scarce.

Not all trials reported on **indeterminate scans**. The NELSON [59,60] and UKLS [63] trials had ,in their work-up algorithms, explicit definitions of indeterminate findings or interval imaging rates. As a consequence the NELSON trial[59,60] LDCT screening result were indeterminate in 10.8% (2,629 out of 24,354) of all scans across three rounds of LDCT screening.

The rates of **additional scan investigations** due to any thorax related anomaly or protocol defined indeterminate scans, varied greatly between trials with a range from 4.5% in the MILD trial [57] to 47.7% in the UKLS trial [63]. **Rates of invasive procedures** per screened individual were low in DLCST (0.5%),

MILD (0.6%), NELSON [59,60] (1.0%) and UKLS [63] (0.6%) but went up to 6.4% in the NLST [6,62], 7.8% in LUSI and 11.4% in the DANTE [43] trials.

Information on the **type of cancer treatment** was generally scarce in trials.

Three trials DANTE [71] MILD [90] and NELSON [92] gave **rates of false negative scans** (defined as scans that remained negative after 12 months follow-up) ranging from ranging from 0.1% NELSON [91], 0.002% MILD [90]bi-annual screening, 0.003% MILD [90] annual screening to 1.3% (DANTE trial [71]).

A simple approximative approach for **overdiagnosis** estimation from a public health perspective, is to calculate the excess incidence, which can be calculated as the difference in cumulative incidence of lung cancer between the LDCT and control arms, and express it as a ratio relative to the total number of lung cancers that are detected in the the LDCT arm. In NLST, the fraction of lung cancers diagnosed during the active screening phase P_a was 11.0% (3.2% to 18.2%). [93] In the NELSON trial the fraction *Pa* at 4.5 years of follow-up and conclusion of the active screening phase was estimated at 19.7% (95% CI, −5.2% to 41.6%) 5.5. years after the closure of the screening program. The overdiagnosis rate was at 11 years of follow-up, however, considerably lower at 8.9% (95%CI -18.2 to 32.4%). [93] [61]

Individuals at high risk for lung cancer (current smoker, former smoker with high tobacco use)					
	Number off participants (RCTs)	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects*	
				Control group risk (only dichotomous outcomes)	Risk difference (dichotomous outcomes) in individuals at high risk of lung cancer
Critical outcomes from randomized comparison					
Lung cancer deaths (direct comparison)	88,006 (7 RCTs)	⊕⊕⊕◯ moderate [a]	**RR 0-80** (0.72 to 0.88)	207 per 10,000	**43 lung cancer deaths less per 10,000** (20 to 58 less per 10,000)
Overall mortality	85,409 (7 RCTs)	⊕⊕◯◯ [b] low	**RR 0.96** (0.92 to 1.00)	878 per 10,000	**36 deaths less per 10,000** (71 less to 0 per 10,000)
Important outcomes from randomized comparison					
Lung cancer stage I or II at diagnosis	85,409 (7 RCTs)	⊕⊕◯◯ [c] low	**RR 2.69** (1.94 – 3.74)	120 stages I –II per 10,000	**202 lung cancers more at early stage I-II detected per 10'000** (112 more to 328 more per 10,000)
Lung cancer stage III or IV at diagnosis	85,409 (7 RCTs)	⊕⊕◯◯ [c] low	**RR 0.79** (0.72 – 0.86)	140 stages III-IV per 10,000	**45 lung cancers less at late stages III-IV per 10,000** (32 less to 67 less)

*For dichotomous outcomes, the risk in the intervention group (and its 95% confidence interval) is based on the control group risk and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio**;**

GRADE Working Group grades of evidence

High quality: We are very confident that the true effect lies close to that of the estimate of the effect

Moderate quality: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different

Low quality: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect

Very low quality: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

Comments. [a] Certainty of evidence downgraded by one level due to moderate risk of selection bias (moderate risk of bias in one trial) [b] Certainty of evidence downgraded by one level due to moderate risk of selection bias (moderate risk of bias in one trial) and one level due to imprecision. Optimal information size not reached. [c] Certainty of evidence downgraded by one level for high risk of performance bias and one level for high risk of detection bias

Table 12 Summary of findings from trials of routine LDCT screening versus no screening or CXR for pooled critical outcomes from randomised comparison (GRADE)

14.Discussion

14.1. Critical outcomes

This HTA report represents an update of the HTA report of Snowsill et al [1] as commissioned and agreed upon with the Cancer Screening Committee. Data from two additional trials (NELSON [59,60]and UKLS [63]) for a total of 7 trials with ≥ 5 years of follow-up could be considered for the analysis of the critical endpoints of lung cancer and overall mortality. Pooled summary estimates indicate a 20% (95% CI 11% - 28%) relative risk reduction from lung cancer mortality of LDCT compared to no screening or CXR, which can be considered as clinically relevant and a 4% (95%CI 0% - 8%) relative risk reduction of overall mortality with little evidence of heterogeneity for both estimates. With an estimated baseline risk of 2% over all trials, this translates into 43 less lung cancer deaths in 10,000 individuals at high risk of lung cancer (95% CI 20 to 58 less lung cancers per 10,000 individuals) over an averaged medium follow-up time of 8.6 years. None of the trials was powered to show a difference in overall mortality. Modelling analyses indicate that a lung cancer screening trial would require 80,000 individuals to show a significant reduction in all-cause mortality (due to a reduction in lung cancer mortality alone) of assumed 2.5% between 11-13 years of follow-up. This reduction in all-cause mortality would only be detectable within a limited period, as individuals' whose lung cancer death is prevented live longer (gaining additional life-years), but will still eventually die of other causes.[101]

Two large trials (NLST [6,62] and NELSON [59,60]) contributed roughly 75% of weight to the pooled estimates with the remaining trials being all underpowered to show a statistically significant effect of LDCT on lung cancer mortality reduction. Certainty of evidence for both mortality outcomes was assessed as moderate and was downgraded by one level for lung cancer mortality due to high risk of attrition bias in one trial and by two levels for overall mortality due to high risk of attrition bias and imprecision in two trials.

As agreed with the purchaser of the report, the large NLST trial, which compared LDCT with CXR, was included into the primary analysis. In a network meta-analysis the pooled estimate of trials directly comparing LDCT with no screening (6 trials) and the indirect comparison of the NSLT trial comparing LDCT with CXR, and 3 trials comparing CXR with no screening was calculated for lung cancer mortality. Pooled estimates from the direct and indirect comparisons of LDCT with no screening were little different confirming the benefical effect of LDCT over no screening. Risk of bias in trials of direct comparisons was low with two trials been rated at moderate risk of bias, wheras two of the three trials examening CXR with no screening were rated at high risk of bias. Risk of bias for indirectness as assessed with the CiNeMA tool was assessed as low, as the network was small with only three nodes,

no relevant heterogeneity or inconsistency was found, and because effect sizes from indirect comparisons were more or less in line with the effect sizes from direct estimates.

While confidence intervals for pooled summary estimates for lung cancer mortality are quite small and the overall risk of bias for mortality outcomes was rated as moderate, one might conclude that we may be quite confident that the true benefit of LDCT screening lies within the estimated margins. The risk of publication bias is likely to be low. No additional unpublished trials were identified similar to other systematic reviews [9] [102] or the HTA report by Snowsill et al [1]. However, it should be kept in mind that 3 trials with nearly 40,000 recruited individuals at high risk of lung cancer from different settings (UK, China, and Japan [69] [9,67]) are still ongoing and results from these trials could modify future pooled effect sizes for lung cancer mortality in either direction towards or off a null effect.

Three recently conducted meta-analyses of RCTs with LDCT screening for lung cancer came up with similar pooled estimates for the benefit of LDCT sceening when compared to no screening or CXR [9] [102] [70]. The US Preventive Services Task Force recently updated its systematic review and recommendation for annual lung cancer screening and "concluded with moderate certainty that annual screening for lung cancer with LDCT has a moderate net benefit in persons at high risk of lung cancer based on age, total cumulative exposure to tobacco smoke, and years since quitting smoking."[103].

In the scope for this HTA report the number of false-positive scans with invasive procedures and the number of false-positive scans with complications from invasive work-up were further defined as critical outcomes. A false positive scan depends on the definition of a positive node or findings in LDCT and work-up algorithms, which largely varied between trials. In addition, two trials (NELSON and UKLS) had a definition of an indeterminated scan and developed work-up procedures for an intermediate repeated scan outside the regular screening visits and work-up algorithms for invasive procedures, which allowed to considerably reduce the number of false positive scans and consequential invasive procedures. In addition, the NELSON trial had a specific volume based definition of a positive lung node that allow to reduce the number of false positive scans in a considerable manner when compared to the remaining trials. The rate of individuals undergoing an invasive procedure due to recall scans or work-ups with non confirmed malignancy was in most trials around or below 5% with one exception, the UKLS trial. In this report invasive procedures are put in relation to the total number of recall scans and not to the number of total screenees because this comparison is from the patient perspective more intuitive and relevant as this figure reflects the risk of an invasive and eventually unnecessary procedure "if something about a LDCT screening scan is wrong".[75] Putting this risk in relation to all screenees is less intuitive to understand for patients and gives very low risk rates as the number of recalls with repeated scans was high in all trials.

The range for recall scans or work-up procedures in relation to the total LDCT screening scans over the entire screening programs was between 1.3% in the MILD trial and 17.0% in the NLST trial. The wide range, again reflects the difference in the definition of positive nodes and work-up procedures. Rates of invasive procedures per screened individual were generally below 1% in several trials (DLCS, MILD, NELSON [59,60] and UKLS [63]) but went up to 6.4% in the NLST [6,62] and 11.4% in the DANTE [43] trials.

In trials reporting data from both groups e.g. the DANTE trial [43] and the NLST [6,62], rates of invasive procedures were generally higher in the LDCT screening arms than in the control groups for both individuals with unconfirmed and confirmed lung cancers.

14.2. Important outcomes

Findings in the two trials (DLCST [89] and UKLS [63]) that evaluated smoking behavior change in relation to lung cancer screening at the broadest study population level did not show that LDCT screening was associated with higher quitting rates when compared to control. However, smoking individuals who quit increased in both trials over time, irrespective of the assigned LCDT screening or no screening. While s smoking cessation counceling should be part of any LCDT screening program, its impact on smoking cessation when combined with the screening program seems not be given.

Psychosocial consequences of LDCT screening were not investigated in all trials and in those trials where these factors were examined, little evidence was found that LDCT screening would increase anxiety, depression or distress.Contrary to expectations, findings in these substudies showed trends that LDCT in longer observation periods was associated with less anxiety.

Overdiagnosis relates to the detection of cancers in screening programs that would have remained clinically inapparent during an individual's lifetime in the absence of screening. Overdiagnosis is driven by indolent tumors or competing causes of death. Data from the NELSON trial indicate the importance of sufficient follow-up to address lead time bias in lung cancer screening programs, as estimates of overdiagnosis rates were at 4.5 years of follow-up - when the active screening phase ended - 19.7% (95% CI, −5.2% to 41.6%) and 8.9% (95%CI -18.2 to 32.4%) after 11 years of follow-up. [59,60] It is of further importance to understand that overdiagnosis due to competing causes of death, play a far more important role in a real world setting. A study comparing real world data from the Surveillance, Epidemiology, and End Results (SEER)-Medicare program, and data from NLST found that 5-year all-cause survival (but not 5-year cancer-specific survival) was significantly worse in the SEER patients with stage I non-small-cell lung cancer who underwent surgery that were cared in routine clinical practice compared to those screenees with same stage cancers undergoing surgery in NLST, and this was especially the case in patients with greater comorbidities. [104]

Ways to control for competing causese of death and to not compromise the efficiency LDCT screening in real world setting are to limit screening in older patients, focus at risk groups with the highest risk

of potential life years lost due to lung cancer, to define and develop prediction tools for such groups, and to develop prediction tools for comorbidities, that may allow to approach individuals most likely to the benefit from lung cancer screening programs.

Ideally in a HTA report, summary estimates of benefits and harms are based on a sound body of evidence from trials investigating a screening intervention and should be put in balance. While pooled estimates for the benefit of lung cancer screening could be derived for the present report, unfortunately, this is not the case for any form of summary estimates for harm that resulted from false positive LDCT scans, and in particular for invasive work-ups that did not confirm the presence of a malignancy. The different definitions of positive lung nodes in trials and in particular the absence of a work-up algorithm in the large NLST trial precluded the generation of summary estimates of harm. The excessive number of follow-up investigations in NLST reflects the strategy chosen by the investigators, to delegate the work-up of abnormal LDCT scans to the patients' private physicians and the fact that not a volume based definition of positive lung nodes was chosen. Thus, from the patients' and resource use perspective NLST does not appear to represent an ideal LDCT screening program for a benefit harm judgement.

For this reason the benefit harm assessment is best practiced by using the NELSON trial, in particular as estimates for lung cancer mortality in NELSON are very close to summary estimates generated in this report (which may also reflect the large weight of NELSON in the overall analysis). Additional arguments are the volume based definition of positive lung nodes in this trial, the used priniciple of indetermine scans for monitoring abnormal LDCT findings and a strict work-up protocol which contributed to much lower false positive rates and invasive work-ups.

In the NELSON trial LDCT arm risk of death from lung cancer was 2.50 deaths per 1,000 person-years and in the control group 3.30 deaths per 1,000 person-years and overall mortality was 13.93 deaths per 1,000 person-years among male participants in the screening group and 13.76 deaths per 1,000 person-years in the control group. During the active screening phase T_0 to T_3 in total 2629 out of 24,354 scans in 7557 participants with a baseline LDCT screening showed a LDCT anormality and 493 participants (6.5%) had to undergo a work-up. In 293 of 493 participants undergoing a work-up, lung cancer was not confirmed (false positive rate of 1.2 % (defined as [recall scans or work-ups – screening detectecd cancers] / total number of individuals) and in 200 individuals cancer was confirmed during the active screening phase of 5.5 years. The risk to undergo an invasive diagnostic work-up procedure over all participants was 1.0%. The estimated overdiagnose rate during the 10 years follow-up period was 8.9%.

To make these figures more intuitive for patients, care providers and health policy decision makers they can be transformed as follows (the figures apply to males only): If 1,000 individuals at high risk of

lung cancer were to be screened at intervals of 1, 2, and additional 2.5 years and then followed for 10 years, 139 individuals would have died from any cause in the screening and 138 would have died without screening but 7 additional individuals (in the best case 12 and in the worst case 2 individuals) in the LDCT group compared to individuals with no screening would have been saved from death from lung cancer. This corresponds to a Number Needed to Screen (NNS) of 134 (95%CI: in the best case 68 and in the worst case 634) to prevent one lung cancer death over 10 years by LDCT screening compared to control. In total 399 of 1,000 individuals were to undergo an additional scan due to an indeterminate finding in regular LDCT scans during the 5.5 years of active screening and 75 would need an invasive diagnostic work up where in 45 individuals lung cancer would not have been confirmed. In 24 of 264 individuals with screen detected lung cancer these 24 cancers would not have become clinically apparent in the absence of screening and were thus overdiagnosed and treated.

14.3. Conclusions

14.3.1. Conclusion – Critical outcomes

LDCT screening in former or current smokers is associated with a 20% relative reduction in lung cancer mortality (moderate evidence) but overall mortality appears not to be affected by LDCT screening (moderate evidence) which may reflect the relative high comorbidity of participants in the examined trials. Risk of abnormal findings particularly at first screening with LDCT is substantial with need of interval scans and diagnostic work-up. False positive scans can be reduced to less than 5% by volume based definitions for lung nodes, strict protocols for interval scans and diagnostic work-up. Overdiagnosis may be an important aspect to consider for implementation programs of LDCT as the at risk population of lung cancer are typically older males at considerable additional risk from comorbidities. Appropriate and well defined criteria for the selection of individuals considered at sufficiently high risk of lung cancer will be paramount for the effectiveness of LDCT screening in a real world setting in addition to a high program adherence. Screening programs for LDCT should use volume based criteria for non-calcified lung node definition, in addition to strict work-up algorithms to limit the number of false positive scans and invasisve work-up procedures.

14.3.2. Conclusion – Important outcomes

Evidence on psychosocial consequences of LDCT is limited as most trials did not investigate anxiety, depression or distress from screening in the entire screened trial population. Where available, data from trial subpopulations did not show any clinically relevant difference of anxiety, depression or distress between LDCT screening and control groups. Likewise no effect of LDCT screening on smoking cessation rates was found.

15.Health economic assessment

15.1. Background

The economic assessment of lung cancer screening consists of a systematic review of the economic literature, a cost-effectiveness analysis, and a budget impact analysis.

The systematic review consisted of an update of the study published by Snowsill et al.[1]. In the cost-effectiveness analysis we updated a previous study with new demographic, epidemiologic, and cost information. Moreover, the effectiveness of LDCT screening was newly based on results of the NELSON trial instead of NLST-based estimates. The budget impact analysis was based on the results of the cost-effectiveness analysis.

15.2. Aims

The aims of the health economic assessment were

- to update the economic literature review by Snowsill et al.[1],
- to assess the cost-effectiveness of LDCT screening in Switzerland,
- to investigate the potential budget impact of LDCT screening in Switzerland.

16.Systematic review of the economic literature

The aim was to update the literature review by Snowsill et al. on the costs and cost-effectiveness of LDCT screening compared to no screening for subjects at high risk for developing lung cancer with emphasis on smokers and former smokers.[1]

16.1. Methods

16.1.1. Literature search strategy

A systematic review of the current economic literature based on the HTA published by Snowsill et al. in 2018 was conducted in Medline (via Ovid), EMBASE (via Ovid), and Web of Science (via Clarivate Analytics) in December 2020. The literature search was conducted by specialists of the Basel University Library. Detailed search strategies are reported in the appendices.

Since the database search by Snowsill et al. was conducted the 5th of January 2017, the updated search was limited to articles published from the end of 2016 onwards.

All types of economic evaluation studies were considered and checked for relevant content: cost-effectiveness analyses, cost-benefit analyses, cost-utility analyses, and cost-minimization analyses.

A non-systematic review of the health economic articles was conducted in Pubmed in October 2021 to identify potentially relevant articles published in 2021.

16.1.2. Screening of the search results

The results of the literature search were screened by two independent reviewers. All articles were screened by title, abstract and, if necessary, by full text review.

In a first step, title and abstracts were screened for relevant quantitative results (e.g., costs, life-years gained (LYG), quality adjusted life-years (QALYs), or incremental cost-effectiveness ratios (ICERs)) or for sentences suggesting potentially relevant content in the full text version.

In a second step, only economic evaluations reflecting the inclusion and exclusion criteria applied by Snowsill et al. and reporting cost per QALY (or LYG) or ICERs were selected.

Only English, French, German, and Italian language studies were considered.

16.1.3. Data extraction, quality assessment, and transferability

For the eligible cost-effectiveness studies (i.e., relevant articles as defined above), data extraction covering the following information were performed:

- Study population (including country, characteristics of included subjects)
- Intervention (e.g., details on screening strategy)
- Comparator(s)
- Setting and perspective of the study
- Cost types included and cost year
- Type of model
- Time horizon
- Discount rate
- Approach to sensitivity analysis
- Incremental cost-effectiveness ratio (ICER)

A brief, qualitative characterization of each relevant study was prepared in the results section, covering methodological approaches taken, main data sources, methodological issues, and potential meaningfulness of the results for Switzerland.

Methodological quality was assessed using the Consensus on Health Economic Criteria (CHEC)-list for economic evaluations as in the HTA conducted by Snowsill et al. [105].

International cost-effectiveness studies were assessed for 'qualitative transferability' to Switzerland. A variety of authors have worked on criteria for assessing such transferability between jurisdictions. Methodologic papers published by O'Brien et al., Welte at al., and Drummond et al. suggested the use

of multistep procedures.[106-109] In the present study, a modified approach based on the above-mentioned procedures is adopted.

The most important criteria for qualitative transferability are already covered by the eligibility criteria. Essentially, for the full-scale health economic evaluation, studies assessing incremental cost-effectiveness must meet the population, intervention, and comparator characteristics of this HTA. Moreover, they have to be performed for countries similar to Switzerland in terms of socioeconomic characteristics.

In short, studies not meeting following CHEC items were regarded as not transferable due to lack of key information:

- CHEC Q1: Is the study population clearly described?
- CHEC Q2: Are competing alternatives clearly described (intervention, comparator)?
- CHEC Q5: Is the chosen time horizon appropriate in order to include relevant costs and consequences?
- CHEC Q6: Is the actual perspective chosen appropriate?
- CHEC Q7: Are all important and relevant costs for each alternative identified?
- CHEC Q9: Are costs valued appropriately (currency, price date, conversion)?
- CHEC Q10: Are all important and relevant outcomes for each alternative identified?
- CHEC Q13: Is an incremental analysis of costs and outcomes of alternatives performed?

16.1.4. Synthesis of finding

Like in Snowsill et al., the characteristics and main results of the included trials- and model-based studies were described.[1] The narrative synthesis provided by Snowsill et al. was updated by including the evidence published after 2016.

16.2. Results

16.2.1. Literature search

A total of 4,811 records were identified from the electronic database searches in December 2020. Following the removal of duplicates (n=1,214), 3,597 citations were screened. Based on titles and abstracts, 3,528 citations were excluded due to inappropriate comparator or non-comparative design, character of a review or commentary piece, or inappropriate outcome measure. A total of 69 citations were included in the full text review. Of these 69 citations, 17 were eligible cost-effectiveness analyses. Fifty studies were excluded due to inappropriate PICO or other reasons (Figure 13). Six additional cost-effectiveness analyses were included from the non-systematic search update conducted in October 2021.

Considering the studies reported in the HTA published by Snowsill et al., a total of 43 cost-effectiveness analyses were included in this report.

16.2.2. Synthesis of characteristics of the identified cost-effectiveness analyses

The characteristics of the included studies are summarized in Table 13. The study publication year ranged from 2001 to 2021. Most of the articles published before 2018 were included in the systematic review conducted by Snowsill et al. in 2017 (n=19).[63,110-128] Only four studies published before 2018 were not identified/included by Snowsill et al. (HTA Ontario 2014, Evans 2016, Cressmann 2017, Treskova 2017). [122] [129-131] A total of 19 articles were published between 2018 and 2021.

One third of the studies were conducted in the USA (n=14), twelve in an European country (5 in the UK, 2 in Germany, 2 in Spain, one in Italy, one in the Netherlands, one in Switzerland). Six studies were conducted in Canada, two in Australia, and three in China. The remaining analyses were conducted in Iran, Israel, Japan, Korea, New Zealand, and Taiwan.

The study population varied considerably across the selected studies. A total of 15 studies included a population reflecting the NLST cohort (persons aged 55–74 years with ≥ 30 pack-years smoking history).[110,111,78,113,115,132,129-131,133-138] The other studies had different age inclusion criteria (starting or stopping age for screening) or smoking intensity criteria. Almost all studies included both males and females smokers. Only two studies included/simulated only male participants (Manser 2004, Whynes 2008).[117-119,121-123,125-127,139,140]

The intervention strategy was also very heterogeneous across studies, ranging from a single LDCT screen to annual, biennial, or triennial LDCT screening from age 50 to 80 years.

In 37 out of 43 studies the comparator was no screening (or no intervention). Five studies compared LDCT with CXR (Tabata 2014, Kumar 2018, Cressmann 2017, Wade 2018, Jaine 2018)[124,129,134,136,137]. The article published by Goffin et al. in 2016 compared biennial LDCT with annual LDCT. [132] Although this comparison did not reflect the PICO of this report, we kept this study as it was also included in the HTA of Snowsill et al.[1]

Most of the included studies used a microsimulation model (n=13), a Markov model (n=12), or a decision tree model (n=12). Other methodological approaches included cohort models (n=4), multistate prediction model (n=1), and a simple deterministic model (n=1). Sensitivity analyses varied across studies, ranging from simple scenario or one-way analyses to probabilistic sensitivity analyses.

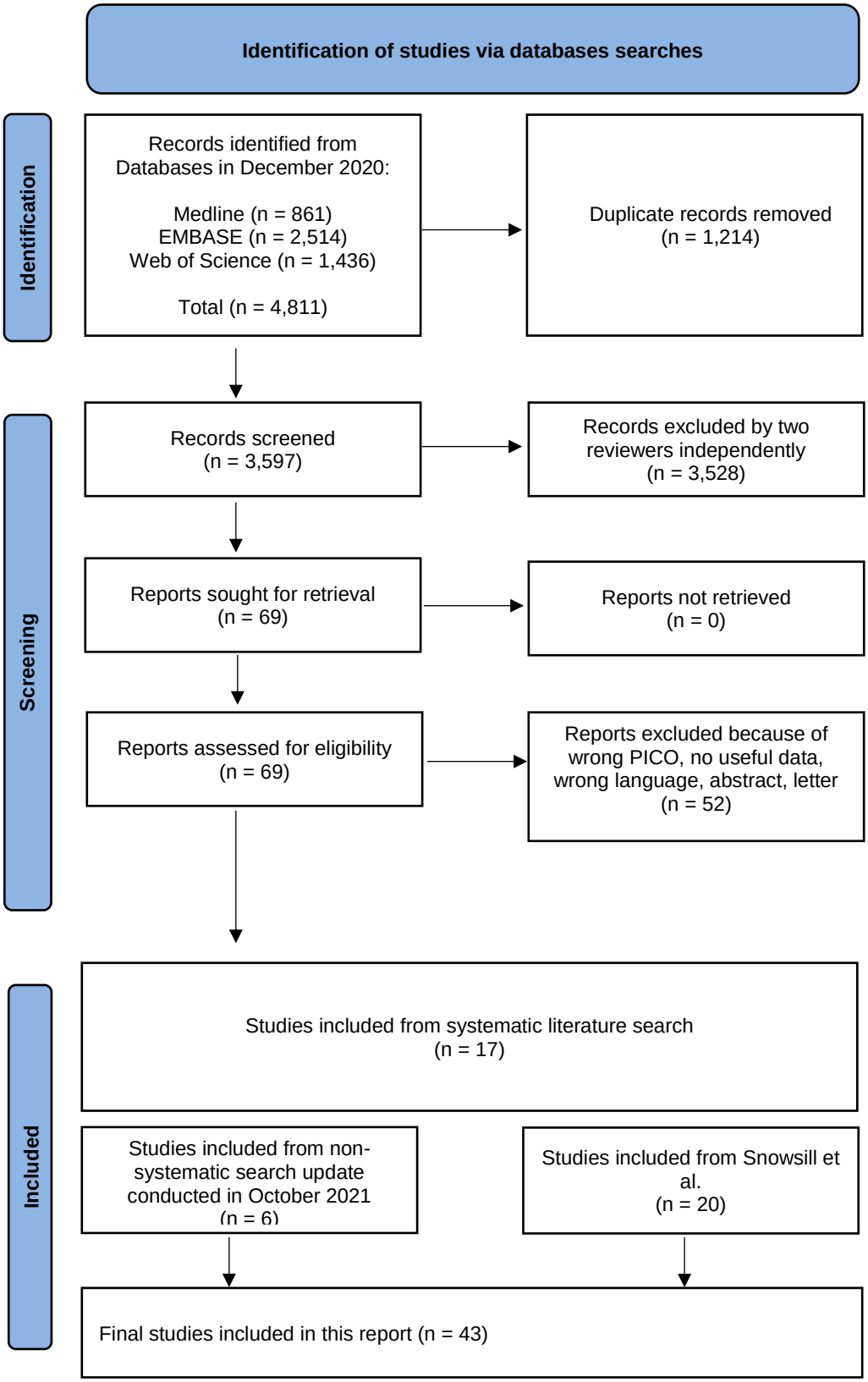


Figure 13 Flowchart health economic literature search

16.2.3. Assessment of quality and transferability

As in the HTA conducted by Snowsill et al., the quality of the selected studies was assessed using the CHEC-list for economic evaluations. [105] For each question of the CHEC-list we judged whether the information was provided (yes), if it was not clearly provided (unclear), or if it was not mentioned (no). To estimate the quality of the included studies we first assigned 1 point for "Yes", 0.5 point for "Unclear", and 0 points for "No". Thereafter, the percentage of the positively answered questions was calculated. The results of the quality assessment for each study are reported in the appendix.

Overall, 80% of the CHEC-questions were answered. However, there were considerable variations across different studies as well as across different CHEC-questions. The study quality ranged from 42% to 100%: Five studies answered less than 60% of the questions, [78,112,113,115,132 118 140 126] 13 studies answered between 60% and 80% of the questions, [110 111 78 113 132 117 139 121 125 127 128 135 141] and 25 studies answered more than 80% of the questions. Although there were few exceptions, studies published more recently tended to have a better quality if compared to older ones.

Concerning the CHEC-questions that were determined as particularly important to consider a study transferable for Switzerland, most items were answered by the majority of the included cost-effectiveness analyses (>90%). Only CHEC Q5 (i.e., "Is the chosen time horizon appropriate in order to include relevant costs and consequences?"), CHEC Q9 ("Are costs valued appropriately? (currency, price date, conversion)") and CHEC Q10 (i.e., "Are all important and relevant outcomes for each alternative identified?") were reported less frequently (77%, 87%, and 83%, respectively).

The biggest issue was the time horizon, which was unclear in four studies (Tabata 2014, Whynes 2008, Esmaeili 2021, Sun 2021 [124 127 142 143]), or was considered too short to capture all relevant costs in other four studies (Goulart 2012, Marshall 2011, Veronesi 2020, Wade 2018 [115 118 137 144]).

Although 26 out of 43 studies had at least one missing/unclear point among the eight CHEC-questions, we considered it fundamental for assessing transferability, and we decided to report characteristics and results for all identified studies.

Study author and publication year	Type of evaluation	Location, price year, currency	Population	Intervention(s)	Comparator(s)	Methodology, Sensitivity analyses
HTA Ontario 2014 [145]	HTA (CUA)	Alberta (Canada), 2012, CAD	NLST cohort (aged 55–74 years with ≥ 30 pack-year smoking history)	Annual and biennial LDCT screening	No screening	Microsimulation model, One-way
HTA Field 2016 # [78]	HTA (CEA and CUA)	UK, 2011–2012, GBP	Adults aged 50–75 years	Risk prediction followed by single LDCT screen	No screening	Decision tree model, One-way and multivariate
HTA Snowsill 2018 [1]	HTA	UK, 2016, GBP	Adult smokers (current or former) aged 55–80	Single, triple, annual, or biennial LDCT screen	No screening	Decision tree model, DSA and PSA
Black 2014 # [111]	CEA and CUA	USA, 2009, USD	NLST cohort (aged 55–74 years with ≥ 30 pack-year smoking history)	Annual LDCT for 3 years, Annual CXR for 3 years	No screening	Decision tree model, One-way
Black 2015 # [110]	CEA and CUA	USA, 2009, USD	NLST cohort (aged 55–74 years with ≥ 30 pack-years smoking history)	Annual LDCT for 3 years, Annual CXR for 3 years	No screening	Decision tree model, One-way
Chirikos 2003 #[112]	CEA	USA, 2000, USD	Adult smokers aged 45–74 years	Annual LDCT for 5 years	No screening	Cohort model, One-way
Cressmann 2017 [129]	CUA	Canada, 2015, CAD	NLST cohort (aged 55–74 years with ≥ 30 pack-years smoking history)	Annual LDCT screening	CXR (assumed to be equal to no screening)	Decision tree model, DSA and PSA
Criss 2019 [146]	CEA	USA, 2018, USD	Current, former, and never-smokers aged 45 years from the 1960 U.S. birth cohort	Annual LDCT according to NLST, CMS, and USPSTF criteria.	No screening	Microsimulation models, Scenario analysis and PSA
Goffin 2015 # [113]	CUA	Canada, 2008, CAD	NLST cohort (aged 55–74 years with ≥ 30 pack-years smoking history)	Annual LDCT for 3 years with smoking cessation	No intervention	Microsimulation model, One-way
Goffin 2016 # [132]	CUA	Canada, 2008, CAD	NLST cohort (aged 55–74 years with ≥ 30 pack-years smoking history)	Biennial LDCT screening for 20 years with/without smoking cessation	Annual LDCT screening for 20 years with/without smoking cessation	Microsimulation model, One-way
Goulart 2012 # [115]	CEA	USA, 2011, USD	NLST cohort (aged 55–74 years with ≥ 30 pack-years smoking history)	LDCT screening (frequency unclear)	No screening	Decision tree model, One-way
Hofer 2018 [147]	CUA	Germany, 2016, EUR	Ever smokers aged 55–75 with ≥ 20 cigarettes per day	Annual LDCT screening	No screening	Markov model, DSA and PSA
Jaine 2018 [134]	CUA	New Zeland, 2011, USD	NLST cohort (aged 55–74 years with ≥ 30 pack-years smoking history)	Biennial LDCT screening	CXR (assumed to be equal to no screening)	Microsimulation model, One-way
Mahadevia 2003 # [116]	CEA	USA, 2001, USD	60-year-old heavy smokers (current and former, > 20 pack years	Annual LDCT to age 80 years)	No screening	Markov model, One-way and multivariate
Manser 2004 # [117]	CEA and CUA	Australia, 2002, AUD	Male current smokers aged 60–64 years	Annual LDCT for 5 years	No screening	Markov model, One-way
Marshall 2001 # [118]	CEA	USA, 1999, USD	General smokers aged 60–74 years	Single LDCT screen	No screening	Decision tree model, One-way
McMahon 2011 # [119]	CUA	USA, 2006, USD	Current and former smokers ≥20 pack-years smoking history	Annual LDCT screening	No intervention	Patient-level microsimulation model,

						One-way
Pyenson 2012 # [121]	CEA	USA, 2012, USD	Current and former smokers aged 50 years with ≥ 30 pack-years smoking history	Annual LDCT from age 50–64 years	No screening	Cohort model, One-way
Pyenson 2014 # [139]	CEA	USA, 2014, USD	Adults aged 55–80 years with ≥ 30 pack-years smoking history	Annual LDCT from age 50–64 years	No screening	Cohort model, One-way
Shmueli 2013 # [123]	CUA	Israel, 2011, USD	Adults aged ≥45 years with ≥10 pack-years smoking history	Single LDCT screen	No screening	Decision tree model, One-way and PSA
Tabata 2014 # [140]	CEA	Japan, NR, JPY	Smokers aged 55–74 years	Annual LDCT	Annual CXR	Decision tree model, One-way and 2-way
Ten Haaf 2017 # [125]	CEA	Canada, 2015, CAD	Adult smokers (current or former) aged 46–75	Eligibility criteria and annual or biennial LDCT screening	No screening	Microsimulation model, One-way
Tomonaga 2018 [148]	CEA	Switzerland, 2015, EUR	Adult smokers (current or former) aged 50–80	Eligibility criteria and annual, biennial, or triennial LDCT screening	No screening	Microsimulation model, One-way
Treskova 2017 [131]	CEA	Germany, NR, EUR	NLST cohort (aged 55–74 years with ≥ 30 pack-years smoking history)	Annual LDCT for 5 years	No screening	Microsimulation model, One-way
Villanti 2013 # [126]	CUA	USA, 2012, USD	Adults aged 50-64 years with ≥ 30 pack-years history	Annual LDCT screening to age 64 years	No screening	Cohort model, One-way
Wade 2018 [137]	CEA and CUA	Australia, 2015, AUD	NLST cohort (aged 55–74 years with ≥ 30 pack-years smoking history)	Annual LDCT for 3 years	CXR (assumed to be equal to no screening)	Markov model (not clearly stated), One-way and PSA
Whynes 2008 # [127]	CUA	UK, 2004, GBP	Men aged 61 years at high risk	Single LDCT screen	No screening	Decision tree model, One-way
Wisnivesky 2003 # [128]	CUA	USA, 2000, USD	Adults aged ≥60 years with ≥10 pack-year smoking history	Single LDCT screen	No screening	Decision tree model, One-way
Yang 2017 [138]	CUA	Taiwan, 2013, USD	NLST cohort (aged 55–74 years with ≥ 30 pack-years smoking history)	Annual LDCT for 3 years	No screening	Markov model (not clearly stated), One-way and PSA
Du 2020 [149]	CEA	Netherlands, 2019 (presumably), EUR	Heavy smokers aged between 20 years old to death	Annual LDCT; for ages 55-80 years for men and ages 50-80 for women	No screening	Microsimulation model, One-way
Hinde 2018 [150]	CEA	United Kingdom, 2015/2016, GBP	Ever-smokers aged 55-74	Annual LDCT screening for 2 years	Not formally stated (assumed to be no LDCT screening programme)	Lung cancer survival curves, Scenario analysis
Evans 2016 [130]	CEA	Canada, 2008 CAD	NLST cohort (aged 55–74 years with ≥ 30 pack-years smoking history)	Annual LDCT screening	No screening	Microsimulation model, Scenario analysis
Kumar 2018 [136]	CEA	USA, 2016 USD	NLST cohort (aged 55–74 years with ≥ 30 pack-years smoking history)	LDCT (3 screenings)	CXR (3 screenings)	Multistate prediction model, One-way, 2-way, 3-way

McLeod 2018 [151]	CEA	New Zealand, NR, NZD	Aged 55-74 years with smoking history (≥40 pack-years)	Biennial LDCT screening	No screening	Markov macrosimulation, One-way
Toumazis 2019 [152]	CEA	USA, 2018, USD	Smokers aged 50-70 years	Biennial LDCT screening	No screening	Microsimulation model, One-way
Veronesi 2020 [144]	CEA	Italy, 2017, (presumably) EUR	Smokers aged 55–79 years (≥ 30 pack-years)	Annual LDCT for 5 years	No screening	Decision tree, DSA and PSA
Esmaeili 2021 [142]	CEA	Iran, NR, IRR	People aged 55–74 (≥ 25 pack-years)	LDCT every 3 years	No screening	Markov cohort simulation, One-way and PSA
Diaz 2021 [133]	CEA	Spain, 2017, EUR	NLST cohort (aged 55–74 years with ≥ 30 pack-years smoking history)	LDCT (@55, @55+60+65, every 2y @55-65, every y @55-65)	No screening	Markov macrosimulation stage-shift model, One-way and 2-way
Gómez-Carballo 2021 [153]	CEA	Spain, 2020, EUR	50–75 years old, current or former smokers (≥ cig./day for ≥25 years or ≥10 cig./day for ≥30 years)	3 LDCT screening rounds at year 1, 2, and 4	No screening	Markov Model, DSA and PSA
Griffin 2020 [154]	CEA	United Kingdom, 2016, GBP	Current or former 60-75 years smokers	Single LDCT screen	No screening	Individual patient simulation, One-way and PSA
Kim 2021 [135]	CUA	Korea, 2015, USD	NLST cohort (aged 55–74 years with ≥ 30 pack-years smoking history)	Annual LDCT screening	No screening	Markov models, One-way
Sun 2021 [143]	CEA	China, 2018, USD	NLST cohort (aged 55–74 years with ≥ 30 pack-year smoking history)	Annual LDCT screening	No screening	Markov model, One-way
Yuan 2021 [141]	CEA	China, 2020, USD	40 years old, heavy smokers in 2020 (≥ 30 pack-years smoking history)	Annual LDCT screening	No screening	Markov model, One-way and PSA

Study included in the HTA published by Snowsill et al. in 2018.
CEA: cost-effectiveness analysis; CMS: Centers for Medicare & Medicaid Services; CUA: cost-utility analysis; CXR: chest x-ray; DSA: deterministic sensitivity analysis; HTA: health technology assessment; LDCT: low-dose computed tomography; NR: not reported; PSA: probabilistic sensitivity analysis.

Table 13 Characteristics of the identified articles

16.2.4. Results of the identified health-economic studies

Summarizing the results of the identified cost-effectiveness/cost-utility analyses is very difficult considering the heterogeneity of the interventions (e.g., single, annual, biennial, triennial LDCT screening), comparators (no screening or CXR), main source of effectiveness assumptions (e.g., NLST, NELSON, ELCAP, etc.), perspective (e.g., healthcare, payer, insurer, societal), or time horizon (from 1 year to lifetime).

Considering that the aim of this report was to update the HTA published in 2018 by Snowsill et al., in this description of the health economic literature, we decided to focus on the articles published between 2018 and 2021. Special importance was given to the studies that used NELSON data as basis for their effectiveness assumptions.

In general, as also stated in the HTA report by Snowsill et al., a common theme in the study results is that LDCT screening is more costly and more effective than no screening or CXR (NB: studies based on NLST generally assumed that CXR was equal to no screening) (Table 14). Although the reported ICERs varied from less than 10,000 USD/EUR/GBP per QALY/LYG (Veronesi 2020, Schmueli 2013, HTA Fields 2016, Wisnivesky 2003, Esmaeili 2021 [78 123 128 142 144]) to more than 100,000 USD/EUR/GBP (Wade 2018, Mahadevia 2003, Jaine 2018, McMahon 2011 [116 119 134 137]), the great majority of the identified studies reported ICERs ranging between 10,000 and 60,000 USD/EUR/GBP/NZD/CAD per QALY/LYG. If we considered only the 22 studies published in the last three/four years (i.e., all recent studies that were not included in Snowsill et al.), only one study reported high ICER estimations (Wade 2018[137]). Wade et al. conducted a cost-effectiveness analysis comparing annual LDCT screening for three years with CXR (assumed to be equal to no screening) in Australian smokers and ex-smokers aged 55-74 years (≥ 30 pack-years). Using a healthcare perspective and a time horizon of 10 years, they estimated ICERs of AUD 138,000/LYG and AUD 233,000/QALY (roughly equivalent to CHF 91,100/LYG and CHF 153,854/QALY). When only current smokers were included, the ICERs decreased to AUD 80,500/LYG and AUD 123,000/QALY (equivalent to CHF 53,150/LYG and CHF81,220). It could be argued whether a 10-year time horizon is enough to correctly capture all possible benefits of lung cancer screening. All other studies published after the HTA conducted by Snowsill et al. had ICERs below USD/EUR/GBP/NZD/CAD 100,000 per QALY/LYG (most of them were below USD/EUR/GBP/NZD/CAD 50,000 per QALY/LYG).

Concerning the sources of effectiveness assumptions and the cost-effectiveness results in their HTA report, Snowsill et al. reported that there is some evidence that studies based on the Early Lung Cancer Action Project (ELCAP) cohort study predict improved cost-effectiveness for screening versus studies based on NLST or lung cancer natural history models. Similarly, we could argue that the cost-effectiveness analyses that based their effectiveness assumptions on the recently published NELSON study also seemed to lead to improved ICERs for LDCT screening. Five articles identified through this

updated systematic review of the economic literature based their effectiveness assumptions on the NELSON trial (Du 2020, Esmaeili 2021, Gómez-Carballo 2021, McLeod 2018, Yuan 2021 [141 142 149 151 153]).

Du et al. conducted a cost-effectiveness analysis for Netherland and including adult heavy smokers (as defined in the NESLON study).[149] Annual LDCT screening was assumed for eligible males aged 55-80 years and females aged 50-80 years. The reported ICERs were EUR 27,600/LYG for males and EUR 21,100/LYG for females (using an insurer perspective and a lifetime horizon).

Esmaeili et al. conducted an analysis for Iran.[142] The population consisted of adult heavy smokers (55-74 years, ≥ 25 pack-years). LDCT screening every three years was compared to no screening. Although the time horizon was not clearly stated, the reported ICER using a healthcare perspective was IRR 98,515,014.14/QALY (≈ EUR 2,000/QALY).

Gómez-Carballo et al. investigated the cost-effectiveness of three LDCT screening rounds (at year 1, 2, and 4) versus no screening for 50-75 years old current or former smokers (15≥ cigarettes/day for ≥25 years or ≥10 cigarettes/day for ≥30 years) in Spain.[153] Using a lifetime horizon and a healthcare perspective they estimated a ICER of EUR 25,854/QALY.

McLeod et al. conducted an analysis for subjects aged 55-74 years with smoking history (≥40 pack-years) in New Zealand[151]. If compared with no screening, biennial LDCT screening led to an ICER of NZD 34,400 per health-adjusted life-years (HALY) (healthcare perspective and lifetime horizon).

Yuan et al. investigated the cost-effectiveness of annual LDCT screening versus no screening in China[141]. Assuming that screening started at 50 years of age and using a time horizon of 40 years (until age 90 years or death), the estimated ICER was USD 12,547/QALY (healthcare perspective).

Many studies emphasized that the screening strategy (e.g., inclusion criteria for lung cancer screening), the cost of LDCT scans, the effectiveness of screening (sensitivity and stage shift leading to lung cancer detection in early stages) and the incidence/prevalence of lung cancer are key factors affecting the cost-effectiveness of screening. Moreover, studies investigating several screening scenarios suggested that age (of the patients as well as for start/stopping the screening) and smoking history seems to play an important role.

16.2.5. Summary of the published literature

The economic literature on lung cancer screening increased dramatically in the last few years, emphasizing the great interest of the scientific community (and of the policy makers) worldwide.

According to this systematic review, the great majority of the published cost-effectiveness analyses concluded that lung cancer screening may be a cost-effective intervention in various countries.

Analyses based on data from the NELSON trial (published in 2020 and conducted in European countries) confirmed the positive results obtained in previous analyses, based on the results of the NLST (based on a US population and published in 2011).

Evidently, the cost-effectiveness of screening is strongly affected by the screening strategy adopted. So far, strategies including subjects at higher risk (based on smoking intensity or PLCOm2012 risk-assessment) seem, to lead to lower ICERs (i.e., to generally better results in terms of costs per LYG or QALY gained). It may be interesting to note that a very recently published analysis comparing the effectiveness of the US Preventive Services Task Force (USPSTF) criteria for lung cancer screening eligibility with PLCOm2012 model eligibility criteria suggested, that patient selection based on a PLCOm2012 risk threshold of at least 1.51% within 6 years appears to be more efficient than USPSTF2013 criteria. [155]

Study author and publication year	Source of effectiveness	Health outcomes	Perspective, time horizon, discount rate	Main results
HTA Ontario 2014 [145]	NLST	QALYs	Payer, lifetime, 3%	LDCT more expensive and more effective than no screening, ICER CAD 92,025/QALY (annual) or CAD 67,396 (biennial)
HTA Field 2016 #[78]	UKLS and estimates of lead time	LYG, QALYs	Healthcare, lifetime, 3.5%	LDCT more expensive and more effective than no screening, ICER GBP 8,466/QALY
HTA Snowsill 2018 [1]	Natural history model calibrated to UKLS and NLST	QALYs	Healthcare, lifetime, 3.5%	LDCT more expensive and more effective than no screening, ICER GBP 28,169-40,034/QALY
Black 2014 # [111]	NLST (assume same outcomes for no screening as CXR)	LYG, QALYs	Societal, lifetime, 3%	LDCT more expensive and more effective than CXR and no screening, ICER (vs. no screening) USD 81,000/QALY. CXR dominated by no screening
Black 2015 # [110]				
Chirikos 2003 # [112]	Hypothetical stage shift	LYG	Payer, 15 years, 7.5%	LDCT more expensive and more effective than no screening, ICER USD 33,557–90,022/LYG depending on achieved stage distribution
Cressmann 2017 [129]	NLST and PanCan	QALYs	Public payer, lifetime, 3%	LDCT more expensive and more effective than CXR, ICER CAD 20,724/QALY
Criss 2019 [146]	NHS/HPFS, SEER, NLST, PLCO, Lung-RADS	LYG, QALYs	Healthcare, 45 years, 3%	LDCT more expensive and more effective than no screening, ICER USD 36,000-51,900/LYG or USD 49,200-96,700/QALY
Goffin 2015 # [113]	Natural history model, partially calibrated to NLST	QALYs	Healthcare, 20 years (lifetime), 3%	LDCT more expensive and more effective than CXR. ICER of triple screen (vs. no screening) CAD 74,000/QALY. ICER of annual screening (vs. no screening) CAD 52,000/QALY. ICER of annual screening vs. triple screen CAD 21,000/QALY (triple screening extendedly dominated)
Goffin 2016 # [132]	Natural history model partially calibrated to NLST	QALYs	Healthcare, lifetime, 3%	Biennial LDCT screening cheaper and less effective than annual LDCT screening ICER of annual vs. biennial ranged from CAD 54,000 to CAD 4.8M/QALY
Goulart 2012 # [115]	NLST	Lung cancer deaths	Healthcare and patient, 1 year, no discounting	LDCT more expensive and more effective than no screening, ICER USD 240,000 per lung cancer death avoided
Hofer 2018 [147]	NLST, German LUSI trial	LYG, QALYs	Payer, 15 years, 3%	LDCT more expensive and more effective than no screening, ICER EUR 19,302/LYG or EUR 30,291/QALY
Jaine 2018 [134]	NLST	QALYs	Healthcare, lifetime, 3%	LDCT more expensive and more effective than CXR, ICER USD 104,000/QALY
Mahadevia 2003 # [116]	Hypothetical stage shift	QALYs	Societal, 40 years (to age 100), 3%	LDCT more expensive and more effective than no screening, ICER USD 116,300/QALY
Manser 2004 # [117]	Diagnostic performance of LDCT based on 'weighted averages of six studies'	LYG, QALYs	Healthcare, 15 years, 3%	LDCT more expensive and more effective than no screening, ICER AUD 57,325/LYG or AUD 105,090/QALY
Marshall 2001 # [118]	ELCAP	LYG	Unclear, 5 years, 3%	LDCT more expensive and more effective than no screening, ICER USD 23,100/LYG. In 'very high-risk' cohort, ICER USD 5,940/LYG

McMahon 2011 # [119]	Natural history model calibrated to tumour registry data and validated against screening studies	QALYs	Societal, lifetime, 3%	LDCT more expensive and more effective than no screening. ICERs for screening consistently above USD100,000/QALY unless positive impact on smoking cessation included
Pyenson 2012 # [121]	ELCAP	LYG	Insurer, 15 years, no discounting	LDCT more expensive and more effective than no screening, ICER USD 18,862/LYG
Pyenson 2014 # [139]	ELCAP	LYG	Insurer, 20 years, no discounting	LDCT more expensive and more effective than no screening, ICER USD 18,452/LYG
Shmueli 2013 # [123]	Single-centre Israeli cohort study	QALYs	Healthcare, lifetime, 3%	LDCT more expensive and more effective than no screening, ICER USD 1,464/QALY
Tabata 2014 # [140]	ALCA, Japanese case–control study	LYG	Unclear, unclear, NR	LDCT more expensive and more effective than CXR, ICERs ranging from JPY 983,000 to JPY 1,942,000/LYG depending on sex and age
Ten Haaf 2017 # [125]	Natural history model calibrated to NLST	LYG	Healthcare, lifetime, 3%	576 screening scenarios evaluated. LDCT screening more expensive and more effective than no screening. 11 screening scenarios and no screening on the efficient frontier. At CAD 50,000/LYG threshold, it is cost-effective to screen annually in 55- to 75-year-olds with ≥ 40 pack-year smoking history (quit ≤ 10 years ago if former smoker), ICER CAD41,136/LYG
Tomonaga 2018 [148]	Natural history model calibrated to NLST	LYG	Healthcare, lifetime, 3%	576 screening scenarios evaluated. LDCT screening more expensive and more effective than no screening. On the efficient frontier 15 of 27 scenarios showed an ICER < EUR 50,000 per LYG.
Treskova 2017 [131]	Natural history model calibrated to tumour registry data and validated against screening studies	LYG	Healthcare, lifetime, 3%	LDCT more expensive and more effective than no screening, ICER EUR 16,754-23,847/LYG
Villanti 2013 # [126]	ELCAP and NLST	QALYs	Payer, 15 years, no discounting	LDCT more expensive and more effective than no screening, ICER USD 28,240/QALY (ELCAP) or USD 47,115/QALY (NLST). Adding smoking cessation nearly doubled QALY gain from screening alone and had lower ICER
Wade 2018 [137]	NLST	LYG, QALYs	Healthcare, 10 years, 5%	LDCT more expensive and more effective than CXR, ICER AUD 138,000/LYG or AUD 233,000/QALY
Whynes 2008 # [127]	ELCAP	QALYs	Unclear (probably healthcare), unclear (perhaps 40 years), 3.5%	LDCT more expensive and more effective than no screening, ICER GBP 13,910/QALY (for men)
Wisnivesky 2003 # [128]	ELCAP	Life-years	Healthcare, unclear, 3%	LDCT more expensive and more effective than no screening, ICER USD 2,500/LYG
Yang 2017 [138]	NLST	QALYs	Public payer, lifetime, 3%	LDCT more expensive and more effective than no screening, ICER USD 19,683/QALY
Du 2020 [149]	NELSON	LYG	Health insurance, lifetime, 4% for costs and 1.5% for LYG	LDCT more expensive and more effective than no screening, ICER EUR 27,600/LYG for men and EUR 21,100/LYG for women

Hinde 2018 [150]	UKLS	QALY	Healthcare, lifetime, 3.5%	LDCT more expensive and more effective than no screening, ICER GBP 10,069/QALY
Evans 2016 [130]	NLST	LYG	Healthcare, 20 years, 3%	LDCT more expensive and more effective than no screening, ICER CAD 42,315/LYG
Kumar 2018 [136]	NLST	QALY	Healthcare, lifetime, 3%	LDCT more expensive and more effective than CXR, ICER USD 60,000/QALY
McLeod 2018 [151]	NELSON	HALY	Healthcare, lifetime, 3%	LDCT more expensive and more effective than no screening, ICER NZD 34,400/HALY
Toumazis 2019 [152]	NLST	QALY	Payer/Insurer, lifetime, 3%	LDCT more expensive and more effective than no screening, ICER USD 46,873/QALY
Veronesi 2020 [144]	NLST	QALY	Taxpayer, 5 years, 3% for costs	LDCT more expensive and more effective than no screening, ICER EUR 3,297/QALY
Esmaeili 2021 [142]	Unclear (probably NLST, I-ELCAP, NELSON)	QALY	Healthcare, unclear, 3.5%	LDCT more expensive and more effective than no screening, ICER IRR 98,515,014.14/QALY
Diaz 2021 [133]	NLST	QALY	Societal, lifetime, 3%	LDCT more expensive and more effective than no screening, ICERs EUR 17,352-34,877/QALY
Gómez-Carballo 2021 [153]	NELSON	QALY	Healthcare, lifetime, 3%	LDCT more expensive and more effective than no screening, ICER EUR 25,854/QALY
Griffin 2020 [154]	NLST, UKLS	QALY	Healthcare, lifetime, 3.5%	LDCT more expensive and more effective than no screening, ICER GBP 28,169/QALY
Kim 2021 [135]	K-LUCAS and the Korean Central Cancer Registry (KCCR)	QALY	Healthcare, lifetime, 5%	LDCT more expensive and more effective than no screening, ICER USD 25,383/QALY
Sun 2021 [143]	NLCSIP, CanSPUC, Chinese studies	QALY	Societal, until age 76, 3% for costs	LDCT more expensive and more effective than no screening, ICER USD 13,473-15,736/QALY
Yuan 2021 [141]	CanSPUC, NLST, NELSON	QALY	Healthcare, until age 90 or death, 3%	LDCT more expensive and more effective than no screening, ICER USD 12,547/QALY (when screening started at 50 years)

Study included in the HTA published by Snowsill et al. in 2018.
ALCA: Anti-Lung Cancer Association; AUD: Australian dollar; CAD: Canadian Dollar; CanSPUC: Cancer Screening Program in Urban China; CXR: Chest X-Ray; ELCAP: Early Lung Cancer Action Project; EUR: Euro; GBP: Great Britain Pound; HALY: health-adjusted life-years; HPFS: Health Professionals Follow-Up Study; HTA: health technology assessment; ICER: incremental cost-effectiveness ratio; IRR: Iranian rial; JPY: Japanese Yen; K-LUCAS: Korean Lung Cancer Screening Project; LDCT: low-dose computed tomography; Lung-RADS: Lung Imaging Reporting and Data System; LYG: Life-year gained; NHS: National Health Service; NLST: National Lung Screening Trial; NR: Not reported; PanCan: Pan-Canadian Early Detection of Lung Cancer; NZD: New Zealand dollar; PLCO: Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial; QALY: quality-adjusted life-years; SEER: Surveillance Epidemiology and End Results; USD: United States dollar; UKLS: UK Lung Cancer Screening Trial.

Table 14 Main results of the identifies studies

17.Cost-effectiveness analysis

The aim of the cost-effectiveness analysis was to investigate the cost-effectiveness of different LDCT screening programs for lung cancer in high-risk populations in Switzerland. The analysis was based on an update of a Swiss cost-effectiveness analysis published in 2018.[148] The specific aim was to update the previous model using effectiveness assumptions based on the results of the NELSON trial (instead of NLST). [61]

17.1. Methods

17.1.1. MISCAN Lung model

Similarly to our previous evaluation, we used the MIcrosimulation SCreening ANalysis (MISCAN) Lung model, a stochastic, microsimulation model. [148] In brief, the model simulates individual life histories in the considered population from birth until death, in the presence or absence of the screening program. Through comparing the life histories in the presence of screening with the corresponding life histories in the absence of screening, MISCAN-Lung can estimate the effectiveness and costs of screening scenarios. MISCAN-Lung was calibrated to individual-level data from the National Lung Screening Trial (NLST) and the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO). [156] [157] While the previous analyses used a version of the MISCAN-Lung model coded in Delphi (Borland Software Corporation, Scotts Valley, California, United States), we have since updated MISCAN-Lung to a new programming structure in Python. Therefore, we have replicated the analyses of our previous investigation to demonstrate the effects of using a new programming structure. Furthermore, we have added analyses incorporating the improved sensitivity found in the Dutch–Belgian lung cancer screening trial (NELSON). The following sections provide additional details on the MISCAN-Lung model.

17.1.2. General model structure

MISCAN-Lung is a semi-Markov model, which generates durations for each state. Individuals are simulated one at a time, which allows future state transitions to depend on past transitions giving the model a "memory". MISCAN-Lung simulates sequences of events by drawing from distributions of probabilities/durations, which makes the results of the model subject to random variation.

MISCAN-Lung consists of several modules: a demography/smoking history generator module, a smoking-dose response module for lung carcinogenesis, a natural history module and a screening module.

17.1.3. Demography/smoking history generator module

First, birth tables, representative for the population under consideration are used to draw a date of birth for each simulated individual. Age, sex and five-year birth-cohort specific smoking initiation probabilities, representative for the population under consideration are used to determine whether an individual initiates smoking and the age of smoking initiation. Upon smoking initiation, persons enter one of five smoking intensity categories. Age, sex, five-year birth-cohort and smoking intensity category specific by averaged number of cigarettes smoked per day are generated for each individual that initiates smoking. If an individual initiates smoking, age, gender and cohort specific smoking cessation probabilities are used to determine whether an individual ceases smoking and the age of smoking cessation. Details on the modeling of the smoking behaviour of the population of Switzerland are presented in the section: “Modelled population and smoking behaviours”.

17.1.4. Smoking related mortality

Upon generating a person’s smoking history, the age of death from causes other than lung cancer is generated, using mortality probabilities based on the person’s smoking history (smoking duration, smoking intensity category and average number of cigarettes per day, smoking status and years since cessation, if applicable), year of birth, age and sex. The maximum age an individual can achieve in MISCAN-Lung is exactly 100 years. Further details on the modelling of the smoking related mortality behaviour of the population of Switzerland are presented in the section: “Modelled population and smoking behaviours”.

17.1.5. Smoking-dose response module for lung carcinogenesis

The smoking-dose response module allows modeling lung carcinogenesis as a function of a person’s age, gender and smoking history. MISCAN-Lung utilizes the two-stage clonal expansion model (TSCE) as described by Heidenreich et al., as its smoking-dose response module (which estimates a person's risk of lung cancer, as a function of age and smoking history). [158] The parameters of the TSCE were obtained through calibration to the Nurses’ Health Study and the Health Professionals Follow-up Study [159]. However, the sex-specific parameters for malignant transformation were recalibrated to data from the National Lung Screening Trial (NLST), the Prostate, Lung, Colorectal and Ovarian cancer screening trial (PLCO) and the Surveillance, Epidemiology, and End Results (SEER) Program. [156] [157]

17.1.6. Natural history module

Lung cancers are assumed to progress sequentially through stages IA to IV, as shown in Figure 14. The probability that a lung cancer progresses to a more advanced preclinical stage or is diagnosed clinically (e.g., diagnosed due to symptoms) is modelled by histology and stage. After clinical diagnosis, lung cancer survival is simulated using sex-, stage-, and histology specific survival estimates, obtained from the Swiss cancer registry for the years 2004-2018.[160] The date of death for individuals with lung cancer is set to the earliest simulated date of death (either due to lung cancer or other causes).

The preclinical durations (in the absence of screening), by histology, stage and gender were calibrated to the rates of screen-detected and interval cancers observed in the NLST and PLCO trials using individual-level data. [156] The preclinical durations (in the absence of screening) are drawn from Weibull distributions. The transition probabilities by histology are reported in Table 15.

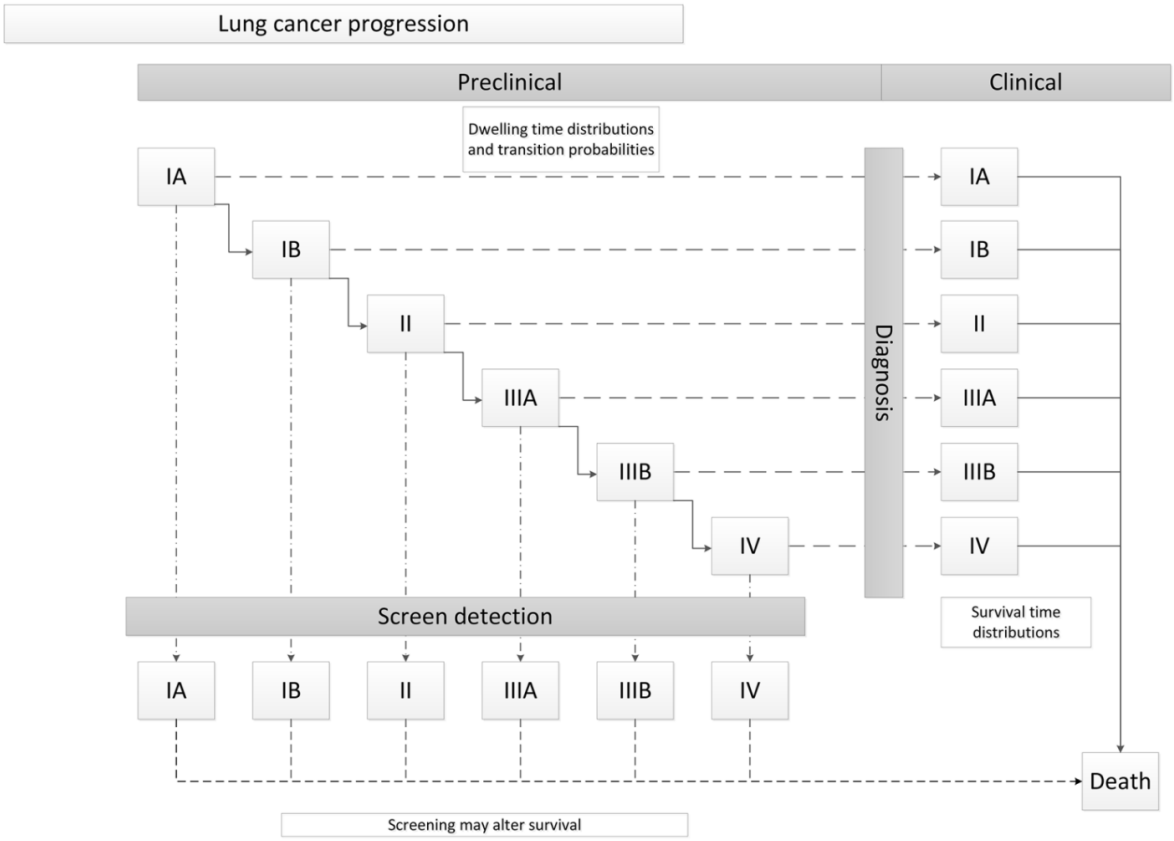


Figure 14 Lung cancer progression in the MISCAN-Lung model

Figure notes: Once lung cancer has developed, it will progress from less advanced to more advanced preclinical stages until it is clinically detected. This process is similar for all histologies, however, the average time spent in the current state differs by histology, preclinical cancer stage and sex. The probability that a cancer progresses to a more advanced preclinical stage or is diagnosed clinically (e.g., diagnosed due to symptoms) is modelled by histology and stage. Screening may detect cancers in each of the preclinical screen-detectable states, depending on the sensitivity of the screening test for the specific histology and preclinical detectable state. Upon detection of lung cancer by screening, a person's life history may be altered. Detection by screening may cure a patient, allowing her/him to resume a normal

(lung cancer free) life history. The probability of cure differs by the stage of detection. After clinical detection or screen detection (without cure) the patient's duration of survival follows a histology and stage specific survival function, which is piecewise uniformly distributed. A person may also die from causes other than lung cancer.

From	To	Adenocarcinoma	Squamous cell carcinoma	Small cell carcinoma	Other non-small cell carcinomas
Preclinical IA	Preclinical IB	0.85	0.87	0.97	0.92
	Clinical detection IA	0.15	0.13	0.03	0.08
Preclinical IB	Preclinical II	0.88	0.85	0.97	0.94
	Clinical detection IB	0.12	0.15	0.03	0.06
Preclinical II	Preclinical IIIA	0.93	0.87	0.97	0.95
	Clinical detection II	0.07	0.13	0.03	0.05
Preclinical IIIA	Preclinical IIIB	0.87	0.81	0.89	0.87
	Clinical detection IIIA	0.13	0.19	0.11	0.13
Preclinical IIIB	Preclinical IV	0.76	0.65	0.80	0.80
	Clinical detection IIIB	0.24	0.35	0.20	0.20
Preclinical IV	Clinical detection IV	1.00	1.00	1.00	1.00
Parameters were estimated by model calibration to individual-level data of the NLST and the PLCO (Adapted from Supplemental material Table S1 of ten Haaf et al, CEBP, 2015)[156]					

Table 15 Transition probabilities by histology

17.1.7. Screening module

Screening may detect cancers in each of the preclinical screen-detectable states, depending on the sensitivity of the screening test for the specific histology and preclinical stage. The model parameters for CT sensitivity by preclinical stage and histology and the effectiveness of CT screening were calibrated to individual-level data from the NLST.[156] Upon detection of lung cancer by screening, a person's life history may be altered. Detection by screening may cure a patient, allowing to resume her or his normal (lung cancer free) life history. The probability of cure differs by the stage at detection. Negative effects of screening, such as overdiagnosis of lung cancer (described subsequently), are also modelled.

To provide an estimate of the effects of incorporating the NELSON trial results in our analyses we increased the CT sensitivity for early-stage cancers (IA, IB and II) across a range of values (+5 percentage points). Table 16 provides an overview of the original sensitivity values and the increased sensitivity values for the early-stage cancers.

17.1.8. Integrating modules

Figure 15 shows an example of how the model integrates the different modules to determine the benefits of screening. The demography/smoking history generator module first generates a date of birth, smoking history and date of death from causes other than lung cancer. This creates a life-history in the absence of lung cancer for Person 1 (shown in life history 1). The smoking-dose response module uses the simulated smoking history to determine whether and when lung carcinogenesis occurs for Person 1 (shown in life history 2). After lung carcinogenesis occurs, the natural history model generates the progression of the cancer, which is diagnosed because of the symptoms in stage II in this example and results in a death due to lung cancer, before the death due to causes other than lung cancer would have occurred (shown in life history 1). In the screening module, a screening examination is simulated, as indicated by the arrow (shown in life history 3). The cancer is detected at the examination and, in this case, the earlier detection allows for successful treatment of the cancer. As a result, the lung cancer death is prevented, and the person's life is prolonged.

Screening may also cause harm, as shown for Patient 2 in Figure 16. In Patient 2 lung cancer also develops, but the cancer would not have been clinically detected without screening (shown in life-history 2). However, the cancer is screen-detected in stage IA during the screening examination simulated in the screening module (shown in life history 3). Thus, in this patient, screening detects a lung cancer that would have never become apparent during the patients' life time if screening had not occurred, resulting in an overdiagnosed case. Thus, for Patient 2 screening does not provide any benefits, but results in life-years with lung cancer care that would not have occurred otherwise (overtreatment).

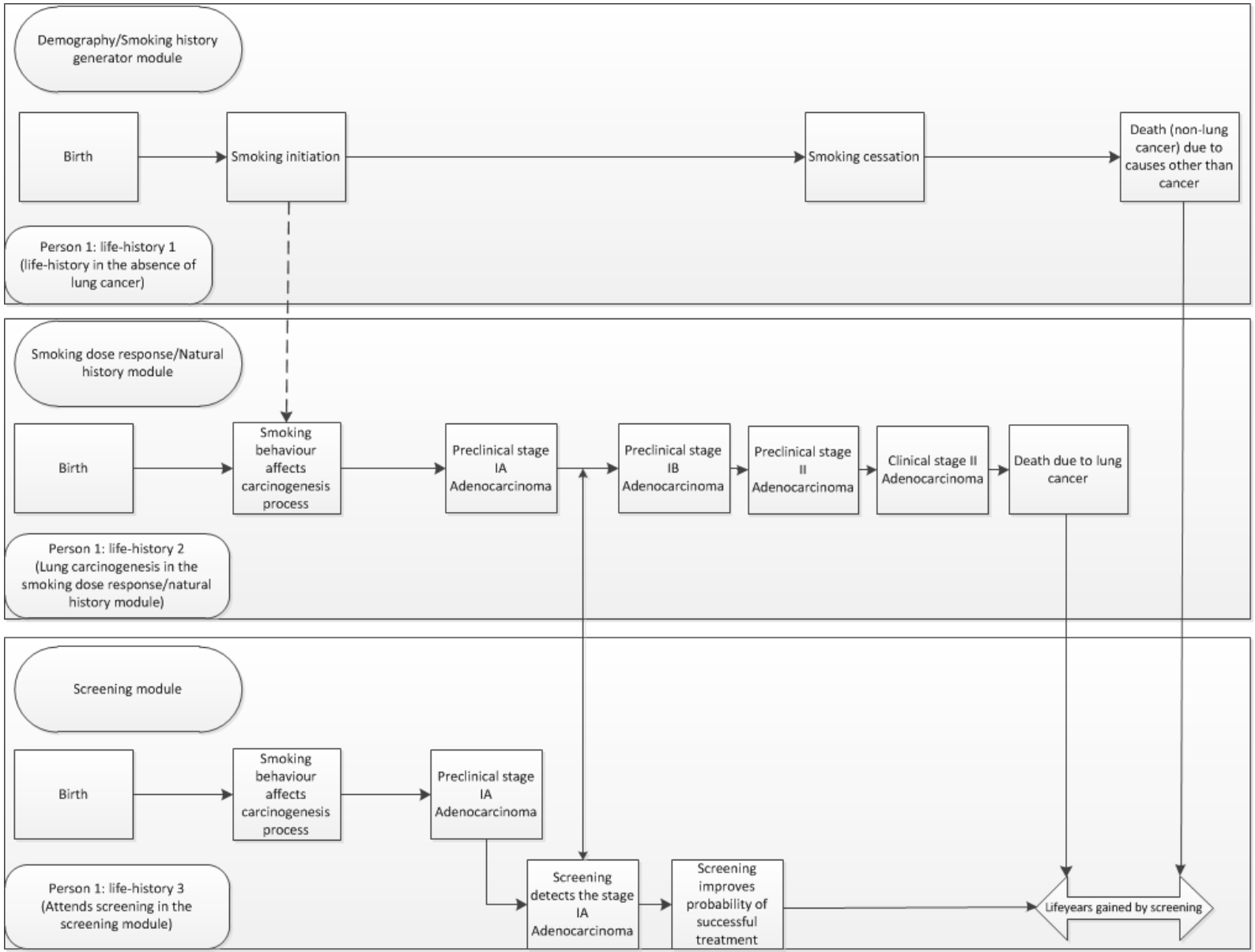


Figure 15 Integrating modules: modelling benefits of screening

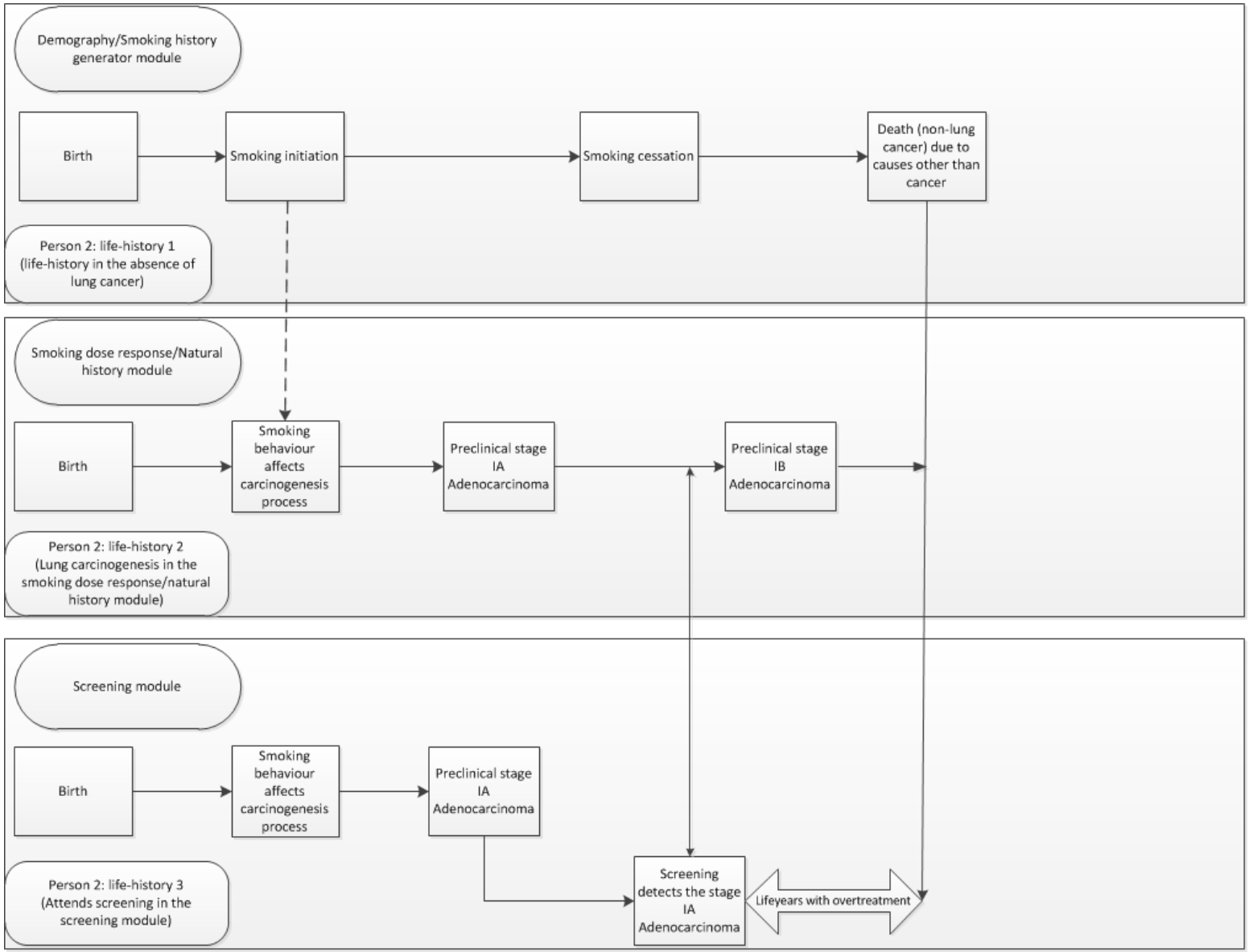


Figure 16 Integrating modules: modelling harms of screening

Histology	Adenocarcinoma			Squamous cell carcinoma			Small cell lung cancer			Other non-small cell lung cancers		
Stage	IA	IB	II	IA	IB	II	IA	IB	II	IA	IB	II
Original sensitivity estimate	56.63%	64.12%	64.48%	30.95%	38.05%	39.19%	8.83%	10.28%	11.19%	20.78%	24.75%	24.78%
5 percentage point increase in sensitivity	61.63%	69.12%	69.48%	35.95%	43.05%	44.19%	13.83%	15.28%	16.19%	25.78%	29.75%	29.78%

Table 16 Sensitivity analyses regarding increased CT sensitivity for early-stage lung cancers

17.1.9. Modelled population and smoking behaviours

In our previous analysis, we evaluated men and women born between 1935 and 1965, incorporating data on smoking behaviour up to 2012. In these analyses we incorporated data from more recent calendar years, as well as expanded our analyses to younger birth-cohorts (e.g., 1965-1969, 1970-1974 and 1975-1979). We included birth-cohorts from 1940 up to and including 1979. The previously included 1935-1939 birth-cohorts are no longer included. These cohorts have passed the age at which they would be eligible for screening at the assumed start of the screening programme of 2023. We set the sizes of the birth cohorts in our model to fit the most recent age- and sex distribution for Switzerland, of 31-12-2020, as provided by the Federal Statistical Office [REF].[161] The following sections detail the methods and assumptions used to generate the smoking behaviours of individuals and the effects of smoking on non-lung cancer mortality in the MISCAN-Lung model for Switzerland.

17.1.10. Smoking initiation

Information on Switzerland-specific current-, former- and never-smoker prevalences up to 2017 was obtained for each birth cohort, by sex from the Swiss Health Surveys conducted by the Swiss Federal Statistical Office.[162] For each five-year birth-cohort, age and sex specific smoking initiation probabilities were estimated, using exponential functions. We assumed ever-smoking prevalence to peak at ages 30-35, after which no further smoking initiation is expected.We therefore matched the smoking intiation rates to the observed ever-smoking prevalence in the 30-35 age group, or the closest age group for which observed data was available. The smoking initiation probabilities were corrected for all-cause mortality using specific all-cause mortality life tables from Switzerland by birth-year, age and sex, obtained from the Swiss Federal Statistical Office. [163] Individuals were assumed to initiate smoking from ages 8 to 29, with age-specific smoking initiation probabilities increasing with age until age 17, after which the probability of smoking initiation decreases, as shown in previous investigations.[164]

17.1.11. Smoking cessation probabilities

Former smokers were defined as smokers who reported to not currently smoke, but had smoked regularly for more than 6 months. For each birth cohort, information on current-, former- and never-smoking prevalences at different ages were used as calibration targets to estimate the age-specific smoking cessation probabilities (the probability that a current smoker at the beginning of that year of age ceases smoking permanently) for each cohort, by sex.

It was assumed that current smokers may cease smoking from age 9 to 100 (the maximum age in MISCAN-Lung) onwards. The probability of successful smoking cessation is assumed to increase with age until age 85, after which the probability of smoking cessation is assumed to be similar to that at age 85. [164] [165] [166] Cessation rates are given by a logistic curve, increasing with age. The parameter for age is estimated for ages 0-30, 30-60 and 60-100 separately, to allow for changes in the interaction between age and smoking cessation. The smoking cessation probabilities by age were estimated for each sex and 5-year birth cohort separately. Cessation rates by this method were found to be lower than previous estimates for Switzerland and the US. To improve the fit of the model to Swiss lung cancer incidence targets, the rates were manually adjusted upwardly by 20%. The smoking cessation probabilities were estimated simultaneously with the mortality probabilities by smoking behaviour (detailed in the “Mortality by smoking behaviour” section).

Table 17 illustrates the observed and estimated ever-smoking prevalence at specific ages for the evaluated Swiss cohorts. Tables 18 and Table 19 show the estimated current-, former- and never-smoking prevalence at different ages compared to the observed prevalence for the 1955-1959 cohort, by sex as an example.

Men			
Cohort	**Age**	**Observed ever-smoking prevalence at this age**	**Estimated ever-smoking prevalence at this age**
1940-1944	50	72.2%	72.3%
1945-1949	45	71.3%	71.3%
1950-1954	40	62.7%	62.6%
1955-1959	35	63.6%	63.8%
1960-1964	35	53.5%	53.4%
1965-1969	35	56.1%	56.1%
1970-1974	35	54.9%	54.9%
1975-1979	35	61.1%	61.1%
Women			
Cohort	**Age**	**Observed ever-smoking prevalence at this age**	**Estimated ever-smoking prevalence at this age**
1940-1944	50	47.4%	47.5%
1945-1949	45	47.6%	47.7%
1950-1954	40	50.4%	50.2%
1955-1959	35	53.1%	53.1%
1960-1964	35	54.9%	54.8%
1965-1969	35	44.7%	44.8%
1970-1974	35	44.0%	44.0%
1974-1979	35	46.5%	46.5%

Table 17 Observed and estimated ever-smoking prevalence at specific ages for the evaluated Swiss cohorts

Age	Observed current-smoking prevalence	Estimated current-smoking prevalence	Observed former-smoking prevalence	Estimated former-smoking prevalence	Observed never-smoking prevalence	Estimated never-smoking prevalence
35	37.1%	34.7%	17.8%	20.2%	45.1%	45.2%
40	34.1%	30.4%	18.1%	24.4%	47.8%	45.2%
45	27.5%	26.7%	21.7%	28.1%	50.8%	45.2%
50	27.2%	23.5%	24.0%	31.4%	48.7%	45.2%
55	25.8%	20.6%	25.3%	34.2%	48.9%	45.2%

Table 18 Observed and estimated current-, former- and never-smoking at different ages for men born between 1960-1964

Age	Observed current-smoking prevalence	Estimated current-smoking prevalence	Observed former-smoking prevalence	Estimated former-smoking prevalence	Observed never-smoking prevalence	Estimated never-smoking prevalence
35	38.8%	36.0%	14.6%	17.5%	46.6%	46.5%
40	38.8%	32.2%	20.5%	21.2%	40.7%	46.5%
45	27.8%	28.8%	21.8%	24.6%	50.4%	46.5%
50	30.7%	25.8%	24.6%	27.7%	44.7%	46.6%
55	29.8%	23.0%	27.0%	30.4%	43.2%	46.6%

Table 19 Observed and estimated current-, former- and never-smoking at different ages for women born between 1960-1964

17.1.12. Cigarettes smoked per day

Data on the average number of cigarettes smoked per day (CPD) were obtained from the Swiss Health Surveys conducted by the Swiss Federal Statistical Office. [162] Furthermore, information on the number of cigars smoked per day (assumed to be equivalent to 2 cigarettes per day), pipes smoked per day (assumed to be equivalent to 2.5 cigarettes per day), and cigarillos smoked per day (assumed to be equivalent to 1 cigarettes per day) was included. Smokers were divided into five smoking-intensity quintiles, ranging from the lightest to heaviest smokers by the reported average number of CPD at each age, similar to Anderson et al.[167] Age-specific values for the average number of CPD per quintile for ages ≥30 were calculated for ages 30, 35, 40, 45, 50, 55, 60, 65 and 70 years, depending on the availability of data for each cohort.

Non-cigarette tobacco products were translated to cigarette equivalents and incorporated in the CPD estimates. With their inclusion the MISCAN-Lung model was found to reproduce lung cancer incidence in Switzerland without additional adjustments to the CPD values, as is sometimes required to compensate for potential underreporting. [168],[169],[170] Linear interpolation was used to fit the average number of CPD by age and quintile between observed CPD values. CPD values were extrapolated by assuming the average CPD value in each quintile decreases by 1% yearly, to reflect the reduction in CPD smoked as individuals age beyond the age of 45-50. [57] Figure 17 and Figure 18 illustrate the variation in average CPD across the five CPD quintiles in the 1960-1964 cohort, for men and women respectively.

In MISCAN-Lung, upon smoking initiation, an individual is randomly assigned to a quintile (with equal probabilities for each quintile) in which the individual will remain until smoking cessation or death. Smoking behaviour was divided into a period of smoking uptake (ages under age 30) and smoking maintenance after the age of 30. The number of CPD for persons younger than 30 are modelled by applying the uptake formulas described by Anderson et al., to the average number of cigarettes per

day at age 30 for that person's smoking quintile:[167]

$$UptakeMale(currentsmokingduration, calenderyear, currentage)$$
$$= -38.578 + 3.342 * \sqrt{currentsmokingduration} - 0.00168$$
$$* \max(79, calenderyear - 1900)^2 - 17.538 * \sqrt{currentage} + 44.967$$
$$* \ln(currentage)$$

$$UptakeFemale(currentsmokingduration, calenderyear, currentage)$$
$$= -56.751 + 0.700 * currentsmokingduration - 0.00163$$
$$* \max(79, calenderyear - 1900)^2 - 3.473 * currentage + 32.8$$
$$* \sqrt{currentage}$$

The uptake formulas are scaled as such, that the number of CPD the person smokes at age 30 matches the one of the average number of CPD in the quintile the person belongs to, regardless of the age of initiation.

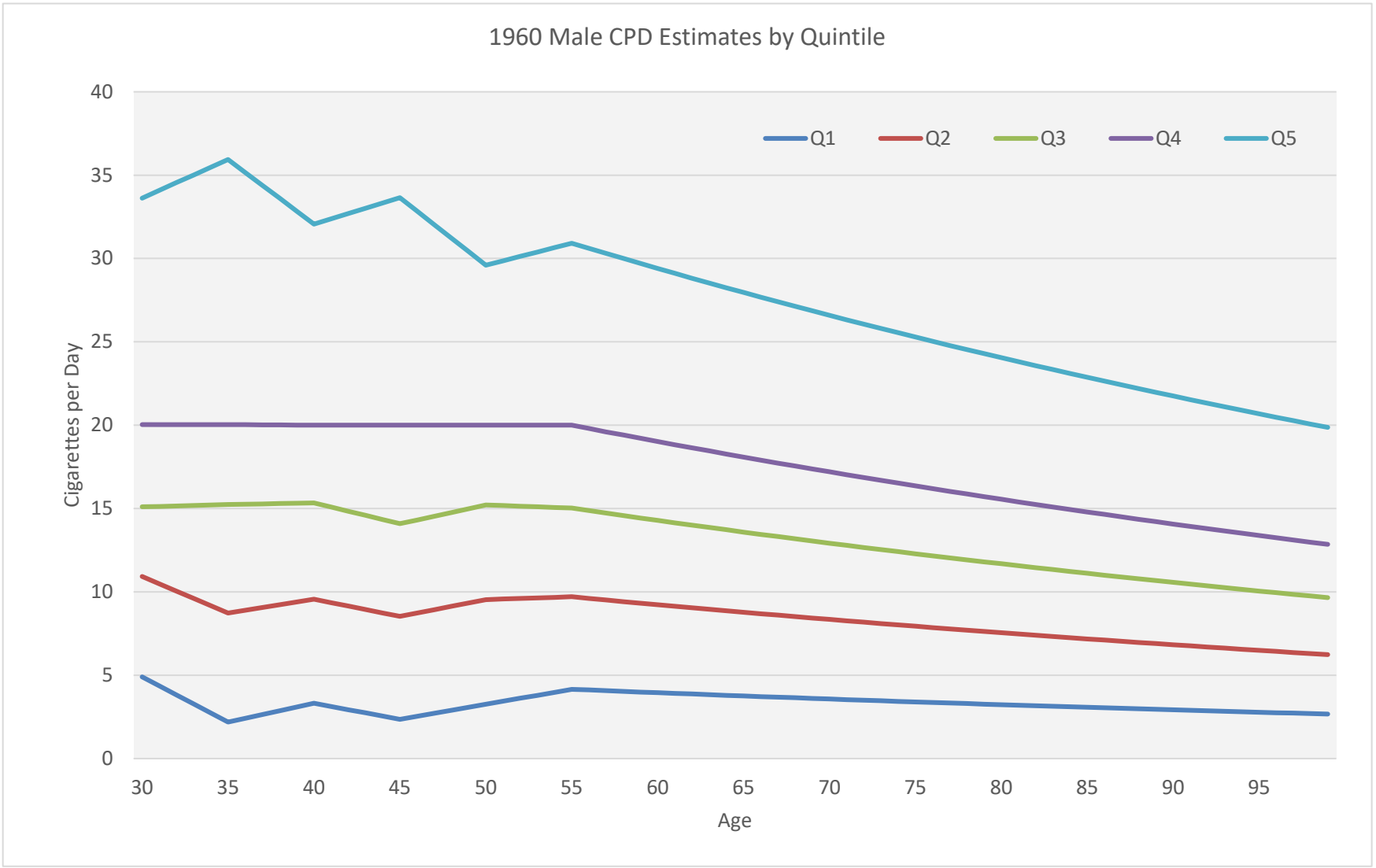


Figure 17 Average number of cigarettes per day for ages over 30, by smoking quintile for men born between 1960-1964

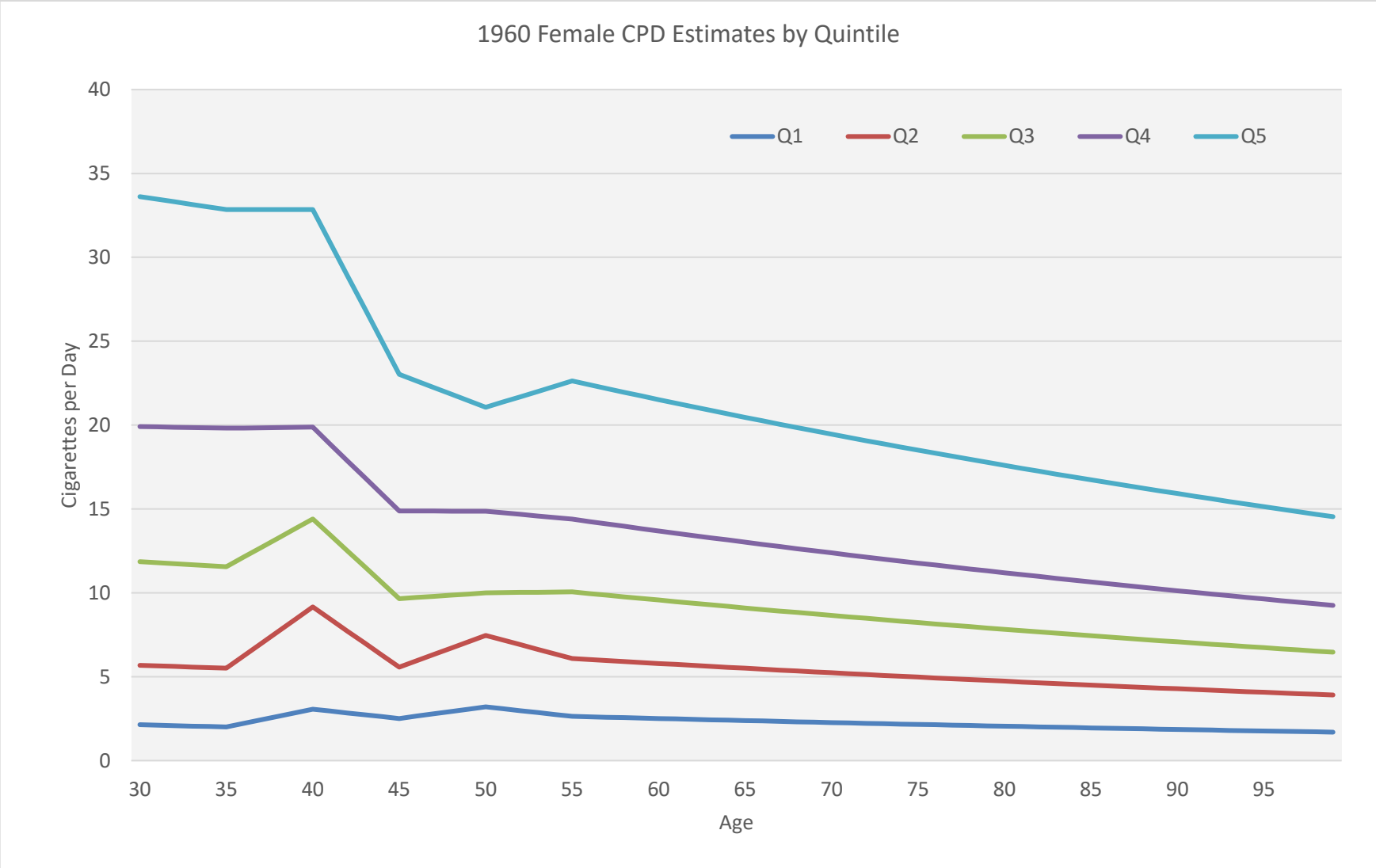


Figure 18 Average number of cigarettes per day for ages over 30, by smoking quintile for women born between 1960-1964

17.1.13. Mortality by smoking behaviour

All-cause mortality life tables specific to Switzerland by birth-year, age and sex, were obtained from the Swiss Federal Statistical Office.[163] To account for competing risks, these life tables were corrected for lung cancer mortality through subtracting the probability of dying from lung cancer from the probability of dying from all causes for each age.[171] Information on lung cancer mortality by five-year age groups and sex was available for years 2004-2018 for ages over 40 and obtained from the Swiss National Institute for Cancer Epidemiology and Registration (NICER).[160] For each available year, the probability of dying from lung cancer was calculated by age group and gender.

Given that mortality rates for lung cancer are low before the age of 40, we assumed the probability of dying from lung cancer to be zero for ages 0-40 for all cohorts. The probabilities of dying from lung cancer for age-groups in the years after 2018 were assumed to be similar to those of the respective age-groups in 2018. The life tables corrected for lung cancer mortality were then further corrected for smoking behaviour, for each birth-year and sex. First, it was assumed that smoking behaviour influences non-lung cancer mortality from age 40 onwards. [171] Thus, before age 40, the non-lung cancer mortality probabilities for never -and ever-smokers with the same birth-year are assumed to be similar:

$$P(nonlungcancermortality_{neversmoker}, currentage, sex, birthyear)$$
$$= P(nonlungcancermortality_{eversmoker,CPDcategory}, currentage, sex, birthyear)$$
$$= P(nonlungcancermortality_{overallpopulation}, currentage, sex, birthyear)$$

From age 40 onward, the non-lung cancer mortality probabilities for never-smokers were assumed to be lower than those of the overall population with the same birth-year, as the overall population includes ever-smokers who have higher non-lung cancer mortality probabilities compared with never-smokers. [171] [172] Therefore, after age 40, the non-lung cancer mortality probabilities of never-smokers were assumed to be similar to the non-lung cancer mortality probabilities of the overall population with the same birth-year, corrected for never-smoking:

$$P(nonlungcancermortality_{neversmoker}, currentage, sex, birthyear) =$$
$$P(nonlungcancermortality_{overallpopulation}, currentage, sex, birthyear) *$$
$$Correctionfactor_{neversmoker}$$

However, as ever-smokers have higher non-lung cancer mortality probabilities compared with never-smokers, the proportion of ever-smokers in the overall population with the same birth-year is expected to decrease at higher ages. Thus, at higher ages, never-smokers will represent a higher proportion of the overall population. As a result, the non-lung cancer mortality probabilities for never-smokers will converge to those of the overall population of that birth-year at higher ages. This convergence is assumed to start from age 70 onward and therefore, the non-lung cancer mortality probabilities of never-smokers from that age onward were assumed to be:

$$P(nonlungcancermortality_{neversmoker}, currentage, sex, birthyear) =$$
$$P\left(nonlungcancermortality_{overallpopulation}, currentage, sex, birthyear\right) *$$
$$\left(Correctionfactor_{neversmoker} + \left(currentage - 69 * \left(\frac{(1 - Correctionfactor_{neversmoker})}{30}\right)\right)\right)$$

As indicated previously, ever-smokers have higher non-lung cancer mortality probabilities compared with never-smokers. However, the non-lung cancer mortality probabilities for ever-smokers are also influenced by the average number of CPD smoked by a person.[171 172] Therefore, it was assumed that non-lung cancer mortality probabilities increased with higher average numbers of CPD. Four categories of average numbers of CPD were defined, similar to Thun et al.: < 10 CPD, 10-19 CPD, 20-39 CPD and ≥ 40 CPD.[172] Furthermore, longer durations of smoking have been indicated to increase non-lung cancer mortality probabilities.[172] Therefore, the increase in non-lung cancer mortality probabilities for current-smokers compared with never-smokers was also assumed to increase with age (as a substitute for smoking duration). In addition, this increase was assumed to differ by smoking quintile, to reflect differences in the average CPD over longer periods of time. Thus, the non-lung cancer mortality probabilities for current smokers from age 40 onward were assumed to be:

$$P\left(nonlungcancermortality_{currentsmoker_{CPDcategory,CPDquintile}}, currentage, sex, birthyear\right)$$
$$= P(nonlungcancermortality_{neversmoker}, currentage, sex, birthyear)$$
$$+ smokingmortalityincrease(age)_{CPDcategory,CPDquintile}$$

Where

$$smokingmortalityincrease(age)_{CPDcategory,CPDquintile} = smokingmortalityincrease_{CPDcategory,CPDquintile} * \left(Smokingagecorrection_{CPDquintile} + ((Currentage - 40) * \left(\frac{\left(1 - Smokingagecorrection_{CPDquintile}\right)}{59} \right)) \right)$$

Previous research indicates that the age of smoking cessation and years since smoking cessation influence the excess risk of mortality due to past smoking behaviour.[171] [172] Overall, the excess risk of mortality decreases for a younger age of smoking cessation and a higher number of years since smoking cessation. Therefore, the excess risk of non-lung cancer mortality was assumed to decrease over time for former smokers:[171]

$$P\left(nonlungcancermortality_{formersmoker_{CPDcategory}}, currentage, sex, birthyear\right) = P(nonlungcancermortality_{neversmoker}, currentage, sex, birthyear) + ((mortalityincrease(age)_{CPDcategory,CPDquintile}) * \exp\left(\left(-0.1711 + (0.00102 * averageCPDoverlifetime) + (0.00171 * QuitAge)\right) * YearsQuit^{1.08}\right))$$

The non-lung cancer mortality probabilities for never- and ever-smokers were estimated simultaneously with the smoking cessation probabilities to match the life tables previously corrected for lung cancer mortality for each birth-year and sex. Figure 20 shows the estimated all-cause mortality probabilities of the overall population for men and women born in 1960 as example. Overall, the estimated all-cause mortality probabilities match those of the observed data. Figure 21 and Figure 22 show the estimated cumulative mortality probabilities of dying from causes other than lung cancer (up to age 85) for never-smokers and current smokers (by smoking quintile) for men and women born in 1962 as examples.

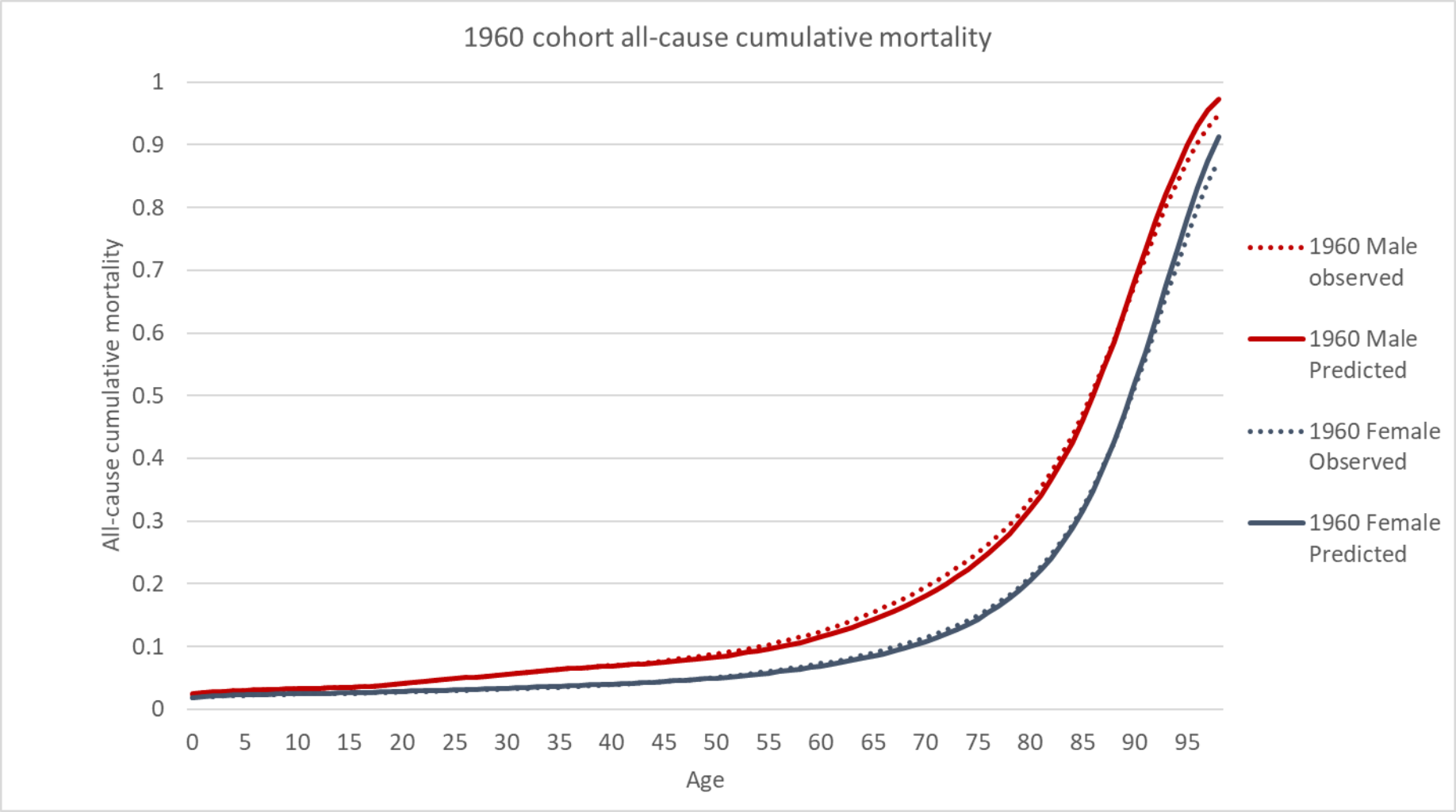


Figure 20 Annual and cumulative probability of dying from all causes for women and men born in 1960

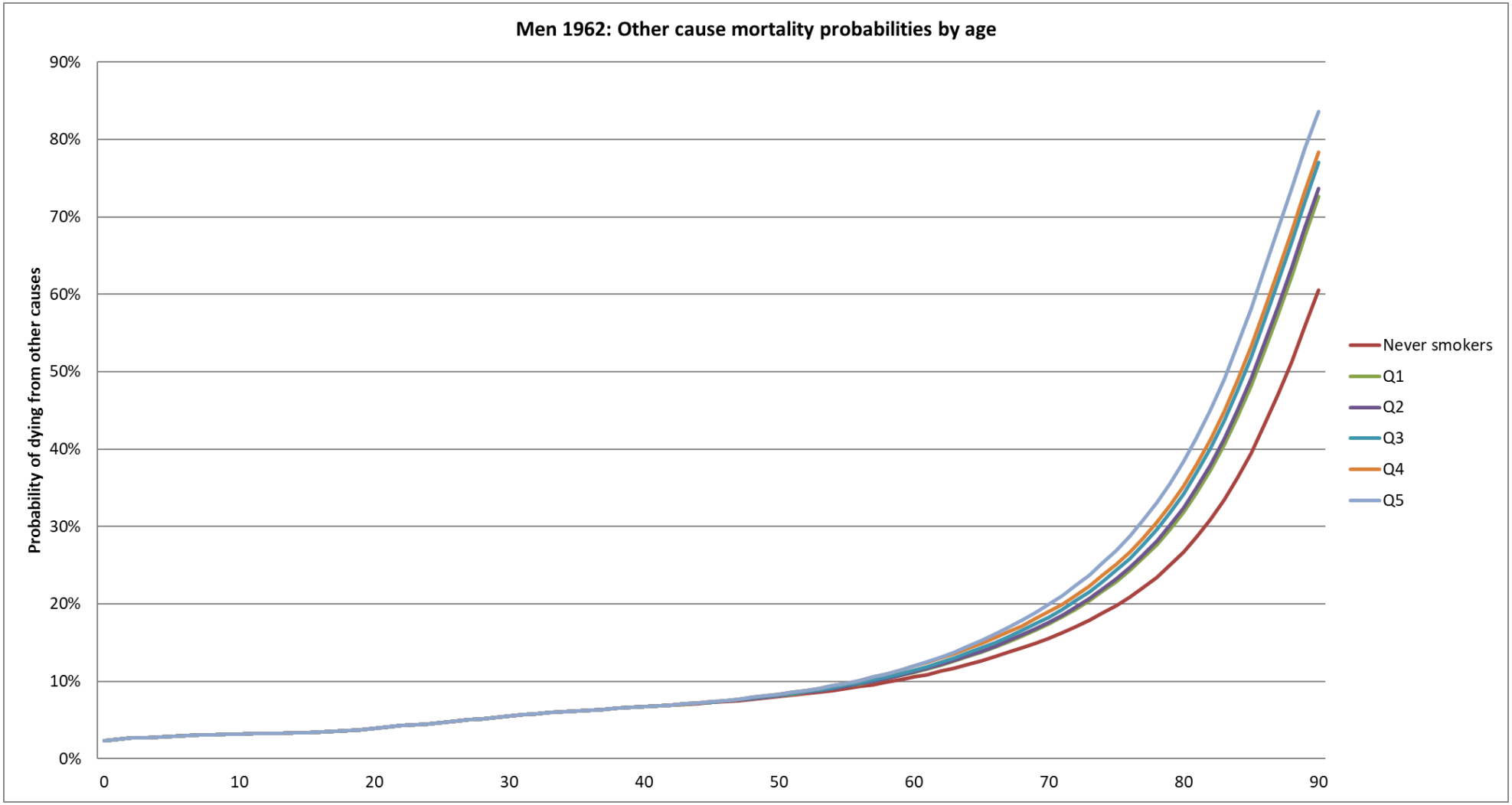


Figure 21 Cumulative probability of dying from causes other than lung cancer for never-smokers and current smokers (by smoking quintile) for men born in 1962

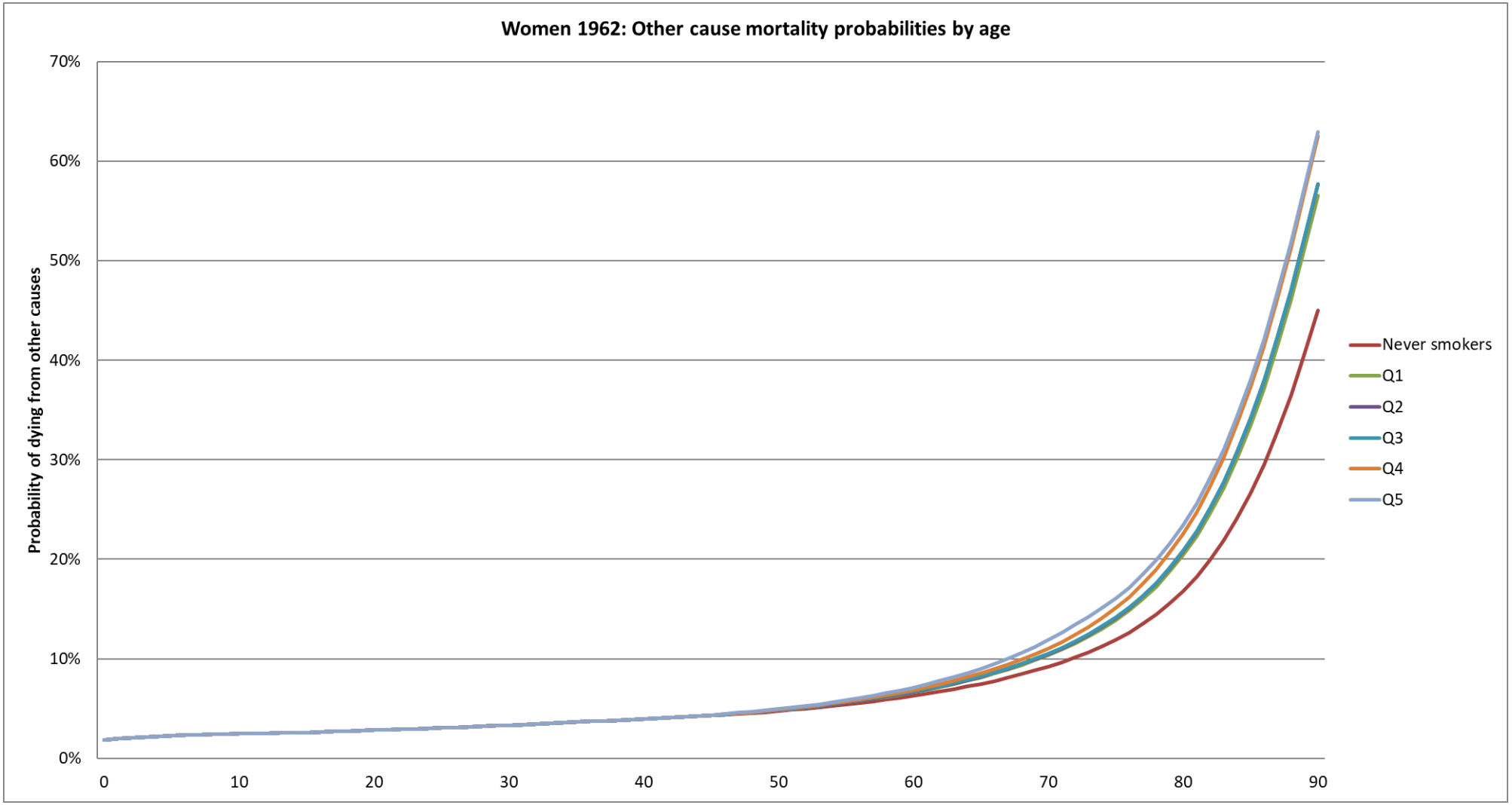


Figure 22 Cumulative probability of dying from causes other than lung cancer for never-smokers and current smokers (by smoking quintile) for women born in 1962

17.1.14. Lung cancer incidence and survival in Switzerland

17.1.15. Lung cancer incidence

Data on the incidence of lung cancer (by age, sex, stage and histology) in Switzerland for years 2004-2018 was obtained from the Swiss National Institute for Cancer Epidemiology and Registration (NICER).[160] Primary, malignant diagnoses (ICD-10 code C33-C34), including multiple lung-cancer diagnoses per person were selected. Data was available from six predominately German-speaking and four predominantly French/Italian speaking Swiss cantons and extrapolated by NICER to reflect the overall population of Switzerland. The available data represents 44% coverage of whole Switzerland.

17.1.16. Lung cancer stages

Lung cancer stage groups were inferred from UICC TNM information according to versions 6 and 7, depending on diagnosis year. The following rules have been applied: (1) if either T, or N descriptors were provided, but M descriptors were missing, the latter were imputed as M0. (2) If only T1 or T2 of TNM-7 was provided, not specifying subclasses a or b, subclass A was imputed.

17.1.17. Lung cancer survival

Five-year survival proportions by age, histology, stage and sex were available from NICER for years 2004-2013, for stages IA, IB, II, III, and IV. Subjects diagnosed at death or with a death certificate as the only source of information (< 3%) were excluded.

17.1.18. Comparison of MISCAN-Lung estimates to observed data

Figure 23 compares the proportions of histological types observed in the Swiss cancer registry data for 2004-2018 by sex, to the proportions estimated by MISCAN-Lung. Overall, MISCAN-Lung reproduces the observed proportions of histological types for both sexes, although a slight decrease in the observed proportion of the histological type `other` may be observed in the newer data for Switzerland.

Figure 24 compares the incidence per 100,000 persons by age group and sex, observed in the Swiss registry data to the incidence estimated by MISCAN-Lung for ages 25-79 (which correspond best to the cohorts born between 1940-1979 which were considered in our analyses). Overall, MISCAN-Lung reproduces the observed incidence well for men, though it somewhat underestimates the incidence at the ages up to age 64, and overestimates the incidence for ages over 70. The same pattern can be seen in replicating the lung cancer incidence among women.

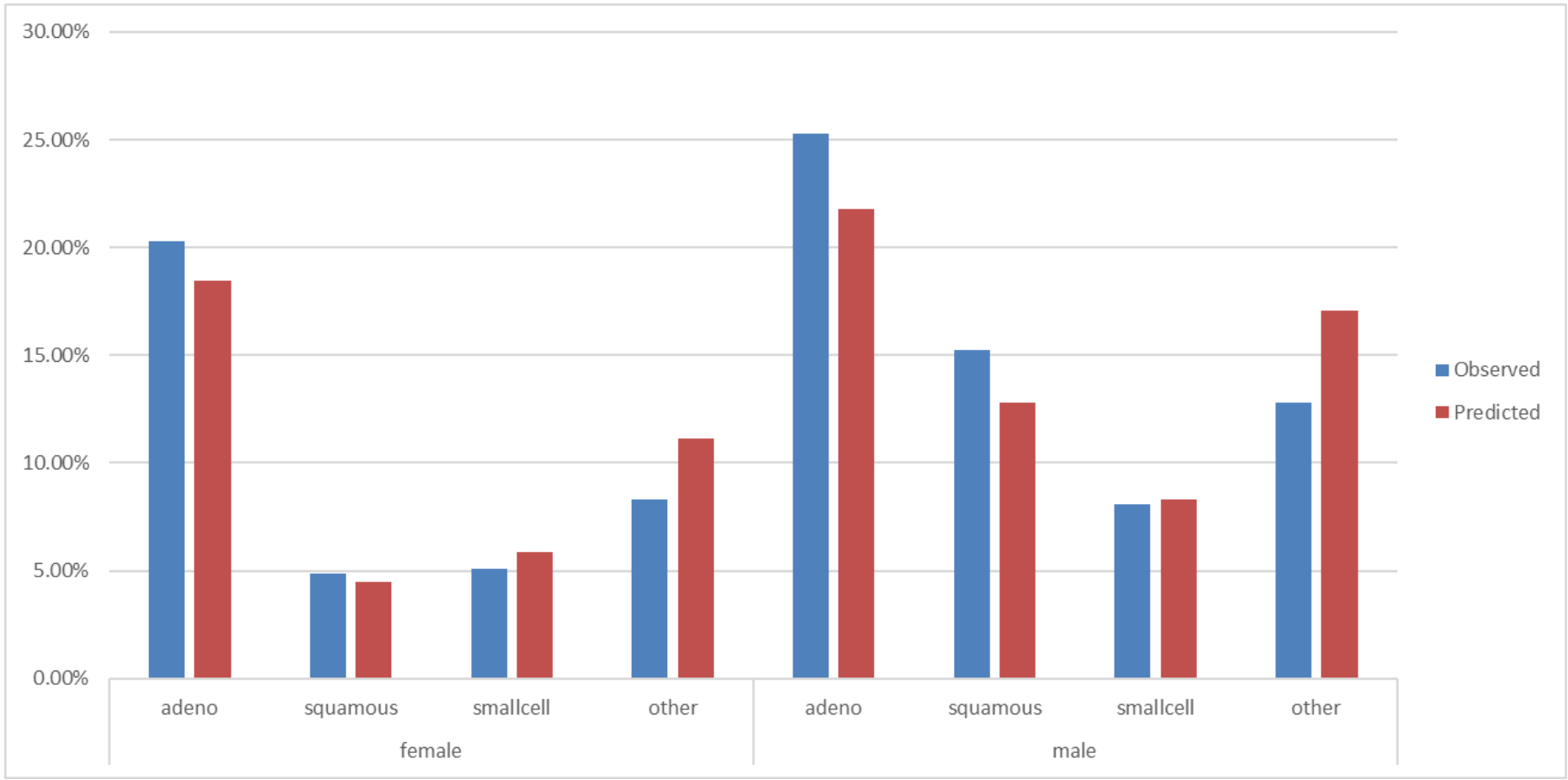


Figure 23 Lung cancer histology distributions estimated by the MISCAN-Lung model compared to the observed lung cancer histology distributions in Switzerland in 2004-2018 for ages 25-79, stratified by sex

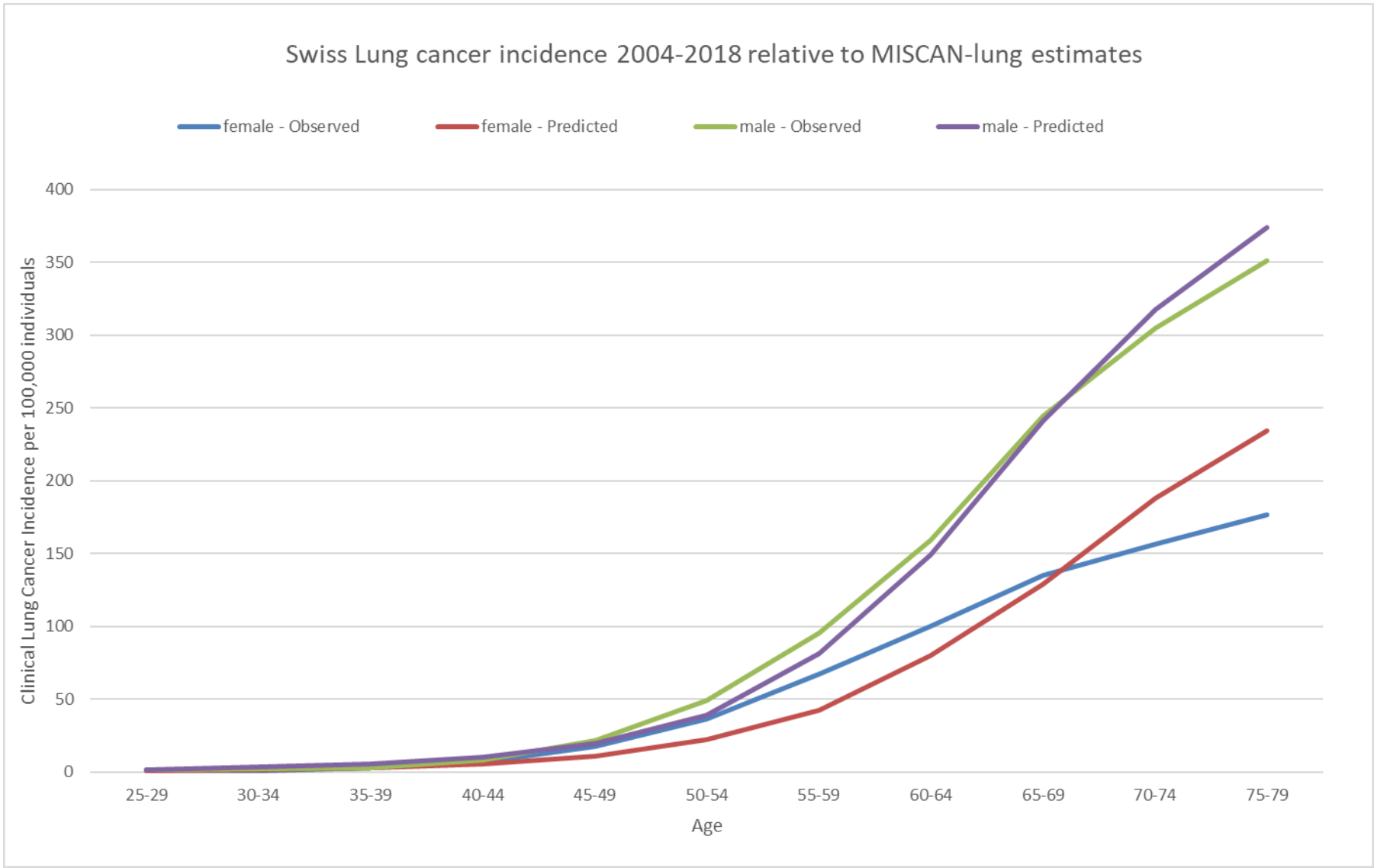


Figure 24 Lung cancer incidence per 100,000 estimated for men aged 25-79 by the MISCAN-Lung model compared to the observed lung cancer incidence in Switzerland in 2004-2018

17.1.19. Modelled scenarios

We replicated the lung cancer screening scenarios modelled in the previous analysis.[173] In addition, we modelled scenarios that considered risk-based selection using the PLCOm2012 model, at different risk thresholds. The PLCOm2012 recently showed superior performance over pack-year based criteria in the prospective International Lung Screening Trial.[155] A truncated version of the PLCOm2012 model was used, which only considers information on age and smoking related eligibility criteria. This truncated version has shown good performance in external validations where it was compared to the full version of the PLCOm2012 model.[174] The truncated version of the PLCOm2012 model uses the same parameter estimates as the original PLCOm2012 model. However, it is assumed that only information on age and smoking history is known. For the analyses in this study, it was assumed that the participant was Caucasian , had a body mass index of 27 kg/m^2(centre value), college education (centre value), no chronic obstructive pulmonary disease, no personal history of cancer, and no family history of lung cancer. Furthermore, we extended the considered range of years since smoking cessation to 25 (as suggested by the 2021 United States Preventive Services Task Force recommendations).[103] A complete overview of the modelled scenarios is provided in Table 20.

Scenario Characteristic	Considered values
Starting age	50, 55 60
Stopping age	75, 80, 85
Screening interval	Annual, Biennial, Triennial
Aggregated smoking criteria	
NLST-like	10, 20, 30, 40 pack-years
NELSON-like	25 y of smoking at least 10 cigarettes per day or 30 y of smoking at least 5 cigarettes per day, 20 y of smoking at least 15 cigarettes per day or 25 y of smoking at least 10 cigarettes per day, 25 y of smoking at least 15 cigarettes per day or 30 y of smoking at least 10 cigarettes per day 30 y of smoking at least 15 cigarettes per day or 35 y of smoking at least 10 cigarettes per day
Maximum years since smoking cessation	10, 15, 20, 25
Risk-based criteria	
PLCOm2012 risk threshold	1.00%, 1.10%, 1.20%, 1.30%, 1.40%, 1.50%, 1.60%, 1.70%, 1.80%, 1.90%, 2.00%, 2.10%, 2.20%, 2.30%, 2.40%, 2.50%, 2.60%, 2.70%, 2.80%, 2.90%, 3.00%, 3.10%, 3.20%

Table 20 Characteristics of the modelled screening scenarios

17.1.20. Considered follow-up rates

The follow-up rates were updated according to those reported for male participants in the **NELSON** trial [61], as shown in Table 21. We assumed that the follow-up rates for the first screening differed from incidence screening rounds. We assumed that the costs and follow-up procedures of true-positive results were included in the phase of care costs of the cancer.

Type of follow-up rate	Baseline screening probability	Incidence screening probability	Consequence
Indeterminate CT result	19.67%	5.08%	Follow-up CT scan
False-positive CT result	1.44%	1.06%	Follow-up CT scan+biopsy

Table 21 Follow-up rates applied in the model

17.1.21. Considered costs

We updated the costs as compared to our previous analyses, as shown in Table 22. All costs were adjusted for inflation. Terminal care costs were additionally inflated by an estimated 85,000 CHF annualized cost to reflect the recent advent of immunotherapy in late-stage lung cancer, based on information from the University Hospital Zurich.

Type of costs	Unit costs in CHF	Occurrence/Maximum duration
Risk-assessment for screening eligibility	80	25% of the population reaching the initial age for screening eligibility
Screening invitation costs	25	Every screening round for eligible individuals
Initial care phase	16,553	Costs per month. From diagnosis to three months.
Continuing care phase	567	Costs per month. From end of initial care to lung cancer terminal care phase.
Terminal care phase	17,885	Costs per month. For the last six months before death.
LDCT screening or follow-up examination	412	Applied to every lc screening, every indeterminate finding and every false positive finding.
Biopsy	1,090	-

Table 22 Considered costs

The analyses were conducted using a healthcare perspective, a lifetime horizon, and a discount rate of 3% (for both costs and effects).

17.1.22. Health state utility values

Since our previous investigation, we have performed a systematic review of health state utility estimates for lung cancer. [175] We have used the values derived from our review to derive estimates for the Quality Adjusted Life-Years gained. The utility estimates corresponding to different health states are shown in Table 23.

Health state	Utility value
Stage I or II lung cancer	0.78
Stage II or IV lung cancer	0.69
Terminal phase	0.59

Table 23 Utility estimates corresponding to different health states

To correctly represent the decrease in utility due to lung cancer it is important to consider the health state utilities of the general population. In fact, assuming that all subjects have perfect health (i.e., a utility value of 1.0) would result in an overestimation of the utility decrement due to lung cancer.

Several studies have shown that utilities in the total population may vary considerably according to country, age, and sex.[176,177,178,179,180]

A Swiss study investigating utilities in the French-speaking part of Switzerland was used to infer general population level health utilities.[180] These values are given in Table 24. The study reports EQ-5D as well as VAS-derived population norm utilities. For our analysis, the EQ-5D values were used.

Age category	Utility value	
	Male	Female
Ages 18-29	0.90	0.86
Ages 30-39	0.87	0.86
Ages 40-49	0.85	0.84
Ages 50-59	0.83	0.81
Ages 60-69	0.83	0.80
Ages 70-79	0.80	0.76
Ages 80-100	0.76	0.74

Table 24 Swiss population norm health utilities by age and sex

17.1.23. Sensitivity analyses

In the first sensitivity analysis, the costs of LDCT and/or lung cancer treatment were varied by 30% for all scenarios on the overall efficiency frontier.

In the second sensitivity analysis we introduced an additional eligibility criterion for all sceening scenarios: In addition to age limitation, smoking intensity criteria, and risk-assessement, only patients with an estimated life expectancy of at least five years were considered eligible. The rationale behind this additional analysis is the ongoing debate on whether patients with limited life-expectancy may or may not benefit from lung cancer screening and, in case of positive findings, of lung cancer treatment. The results of a recently published study already suggested that "*limiting screening to only those with more than 5 years of life expectancy did not greatly affect the resulting estimated benefits (deaths averted or life-years gained) but was estimated to result in fewer harms and considerably fewer overdiagnosed cases.*" [181] According to the authors, this finding was particularly true for screening strategies at older ages. Another comparative modelling analysis evaluating the long-term benefits and harms of risk-based screening compared with current USPSTF recommendations concluded that accounting for limited life expectancy yielded greater improvements in efficiency for risk-based strategies than the USPSTF criteria did. [182]

In the third sensitivity analysis, the effect of discounting was investigated. The discount rate, considered to be equal for both costs and benefits, for the scenario on the efficiency frontier were varied between 0% and 6% in increments of 1.5%.

17.2. Results

The present analyses represent an update of our previous study on lung cancer screening in Switzerland. [148] The main change consists in the effectiveness assumptions, which were previously based on NLST. In this new analyses we implemented effectiveness assumptions based on the NELSON trial. In particular, an increase in LDCT sensitivity for stages IA, IB, and II as well as a decrease in false-positive rates were implemented.

In addition to the effectiveness assumptions, demographic, epidemiologic, smoking behaviour, and cost input parameters were updated with the most recent information available. In comparison with our previous study, the model newly included utility values for the general Swiss population as well as for subjects with lung cancer. This allowed to calculate ICERs as costs per QALY gained (instead of cost per LYG). Another important change concerns the costs. Besides the costs input used in the previous study, updated to 2021 according to the Swiss inflation rate, we newly included costs for patient invitation to screening, risk assessment, and immunotherapy.

17.2.1. Comparison of original results based on NLST effectiveness and new estimations based on NELSON effectiveness

To compare the previously published analyses based on NLST effectiveness ("original") with the new estimations based on NELSON effectiveness, a total of 2,972 scenarios were modelled.

For both effectiveness assumption we modelled one no screening scenario, 432 NLST-like scenarios (i.e., scenarios based on NLST inclusion criteria), 432 NELSON-like scenarios (i.e., scenarios based on NELSON inclusion criteria), and 621 risk-based scenarios were modelled (i.e., scenarios based on PLCOm2012 risk assessment).

The costs and LYG of all screening scenarios versus no screening are illustrated in Figure 25. In general, it can be noted that scenarios based on NELSON effectiveness led to more LYG and lower costs if compared to the original scenarios based on NLST effectiveness. The average cost-effectiveness ratios (ACERs) comparing each scenario with no screening for the models based on NLST-effectiveness ranged from CHF 27,392/LYG to CHF 64,281/LYG, whereas model based on NELSON effectiveness led to ACERs ranging between CHF 12,038/LYG to CHF 31,214/LYG (Of note: A lower ACERs suggest an improvement in cost-effectiveness). Compared with the ACERs using NLST effectiveness assumptions, the ACERs based on NELSON assumptions were between 17.6% and 74.4% lower.

Figure 26 illustrates the efficiency frontiers of the scenarios based on NLST effectiveness (original) and NELSON effectiveness. It can be noticed that the assumption of an increased LDCT sensitivity based on the NELSON effectiveness led to a shift of the efficiency frontier. The new analyses based on NELSON suggest that lung cancer screening may be more efficient and less expensive than previously estimated using NLST effectiveness assumptions. It should be emphasized that the changes in costs is presumably due to the fact that the new calculations newly included the costs for immunotherapy, which are particularly high. Immunotherapy is mainly given to patients with advanced lung cancer stages (stages III and IV). Since screening generally lead to a stage shift (i.e. patients are diagnosed at earlier stages), less patients require immunotherapy.

17.2.2. Characteristics of the scenarios on the new efficiency frontier

Table 25 summarizes the characteristics and cost-effectiveness of the scenarios on the efficiency frontier for the scenarios based on NELSON assumptions (including NLST-like, NELSON-like, and risk-based scenarios). Compared to no screening, the scenario on the efficiency frontier cost between CHF 14,452 and CHF 37,959 per QALY gained. Almost all scenarios on the efficiency frontier (19 out of 21) had a ICER below CHF 100,000 per QALY gained.

In contrast to our previous analyses in which the frontier started with triennial sceening for heavy smoker, the frontier based on the new analyses starts with biennial screening. Continuing on the efficiency frontier, annual scenarios with broader inclusion criteria (i.e., larger screening age interval, lower smoking intensity, and lower risk) were modelled. One main difference with the previous analyses is the introduction of risk-based strategies that appear to dominate a large amount of NLST-like and NELSON-like scenarios.

Table 26 provides an overview of selected outcomes for scenarios on the overall efficiency frontier. Depending on the inclusion criteria, the percentage of the smoking population that would be screened ranged between 6.8% and 27.0%. The total number of CT screens, CT scans (i.e., follow-up CTs), biopsies, false-positive scans, and overdiagnosed cases differed considerably across screening strategies (by a factor higher than 10).

Concerning lung cancer incidence and lung cancer-related mortality, the model assuming no screening estimated the detection of 6,784 lung cancer cases and a total of 4,674 lung cancer deaths (in a population of 100,000 persons). The introduction of lung cancer screening would lead to a higher number of detected lung cancer cases and a lower number of cancer deaths. Depending on the scenario adopted, the number of cases would range between 6,799 (+15 cases per 100,000 persons compared to no screening) and 6,981, (+197 cases per 100,000 persons compared to no screening), while the number of lung cancer deaths would range between 4,471 (-4.3%) and 3,593 (-23.1%). In other words, the number of lung cancer deaths prevented with screening would range between 203 per 100,000 persons for the first scenario on the efficiency frontier to 1,071 per 100,000 persons for the last scenario on the efficiency frontier.

In our previous study the number of false positive screens per 100,000 persons (based on NLST effectiveness) were particular high, ranging between 7,651 and 63,435. The new analyses based on NELSON false-positive rates showed a drastic decrease, with false positive screens ranging between 360 and 8,290 per 100,000 persons. Although less pronounced, the number of overdiagnosed lung cancers decreased (from 67 to 338 per 100,000 persons in our previous study to 15 to 196 per 100, 000 persons in the current analyses).

Depending on the screening scenario, the number of individuals needed to screen per LYG would range between 2 and 3 (i.e. one needed to screen 2-3 persons at risk to gain one life-year), while the number of individuals needed to screen per death avoided would range between 21 and 41. The number of LDCT screen per lung cancer death avoided would range between 155 and 434 LDCT screens per LYG.

Considering that the investigated birth cohorts 1940–1979 (aged 41–80 years in 2020), represented approximately 4.1 million individuals in Switzerland in 2020, the impact that lung cancer screening might have if the whole population would be considered is obviously much higher (to get an idea of the overall impact assuming that all eligible persons would participate, the numbers provided in the table should be multiplied by 41).

For example, assuming all eligible persons would participate, an annual screening for 55-75 years old persons based on a PLCOm2012 risk assessment threshold of 2.3% (scenario name: plco_55_75_0.023_1) would lead to the prevention of 17,022 lung cancer deaths.

The following chapters briefly illustrate the characteristics and selected outcomes for the efficiency frontiers using only NLST-like inclusion criteria (Table 27 and Table 28), only NELSON-like inclusion criteria (Table 29 and Table 30), or risk-based inclusion criteria (Table 31 and Table 32). In general, it is worth mentioning that the efficiency frontiers of the three inclusion criteria are close to each other (Figure 27). As it can already be deducted from the overall efficiency frontier, scenarios based on the PLCOm2012 risk assessment seem to be slightly more efficient than those focusing on smoking intensity criteria (i.e. using NELSON-like and NLST-like inclusion criteria).

17.2.3. Scenarios based on NLST inclusion criteria

Compared to no screening, the scenarios on the efficiency frontier had ACERs ranging between CHF 14,452 and CHF 37,959 per QALY gained (Table 27). The most cost-effective scenarios were biennial screening for patients aged 60 to 75 years with particularly high smoking intensity (30-40 pack-years). Continuing on the efficiency frontier, annual screenings for larger age groups (up to 50 to 85 years) and lower smoking intensity (i.e. 10 pack-years, with up to 25 years since smoking cessation) appears. If compared between each other, almost all screening scenarios had an ICER below CHF 100,000 per QALY gained. Only the last scenario on the efficiency frontier was higher (CHF 128,923 per QALY gained).

The percentage of the screened population ranged between 6.8% and 27.0%, while the number of lung cancer death avoided increased from 203 per 100,000 persons in the first scenario on the efficiency frontier to 1,071 per 100,000 persons in the last scenario (Table 28).

17.2.4. Scenarios based on NELSON inclusion criteria

As in the scenarios based on the NLST inclusion criteria, the first scenario on the efficiency frontier based on NELSON criteria assumed biennial screening for a rather restricted age group (from 60 to 78/80 years) of heavy smokers (e.g. 30 years x 15 cig/day or 35 years x 10 cig/day) (Table 29). The

ACERs ranged from CHF 16,236 to CHF 33,347 per QALY gained, while all ICERs were far below CHF 100,000 per QALY gained (range CHF 16,236 – CHF 61,008 per QALY gained).

The percentage of the population screened ranged between 10.0% and 21.9%, while the number of lung cancer deaths avoided increased from 263 at the beginning of the efficiency frontier and 1,044 at the end of it (Table 30).

17.2.5. Scenarios based on PLCOm2012 risk assessment threshold

The first nine scenarios on the efficiency frontier based on the PLCOm2012 risk assessment included first biennial and then annual screening for patients aged 60 to 75 years and a risk threshold ranging between 2.5% and 3.2% (Table 31). Following scenarios on the efficiency frontier extended the age range (up to 50-85 years) while decreasing the risk threshold (down to 1.1%). In general, ACERs as well as ICERs were geneally lower if compared to the scenarios based on NLST or NELSON inclusion criteria: ACERs ranged between CHF 12,111 and CHF 25,231 per QALY gained, while ICERs raged between CHF 14,595 and CHF 30,889 per QALY gained (Table 32).

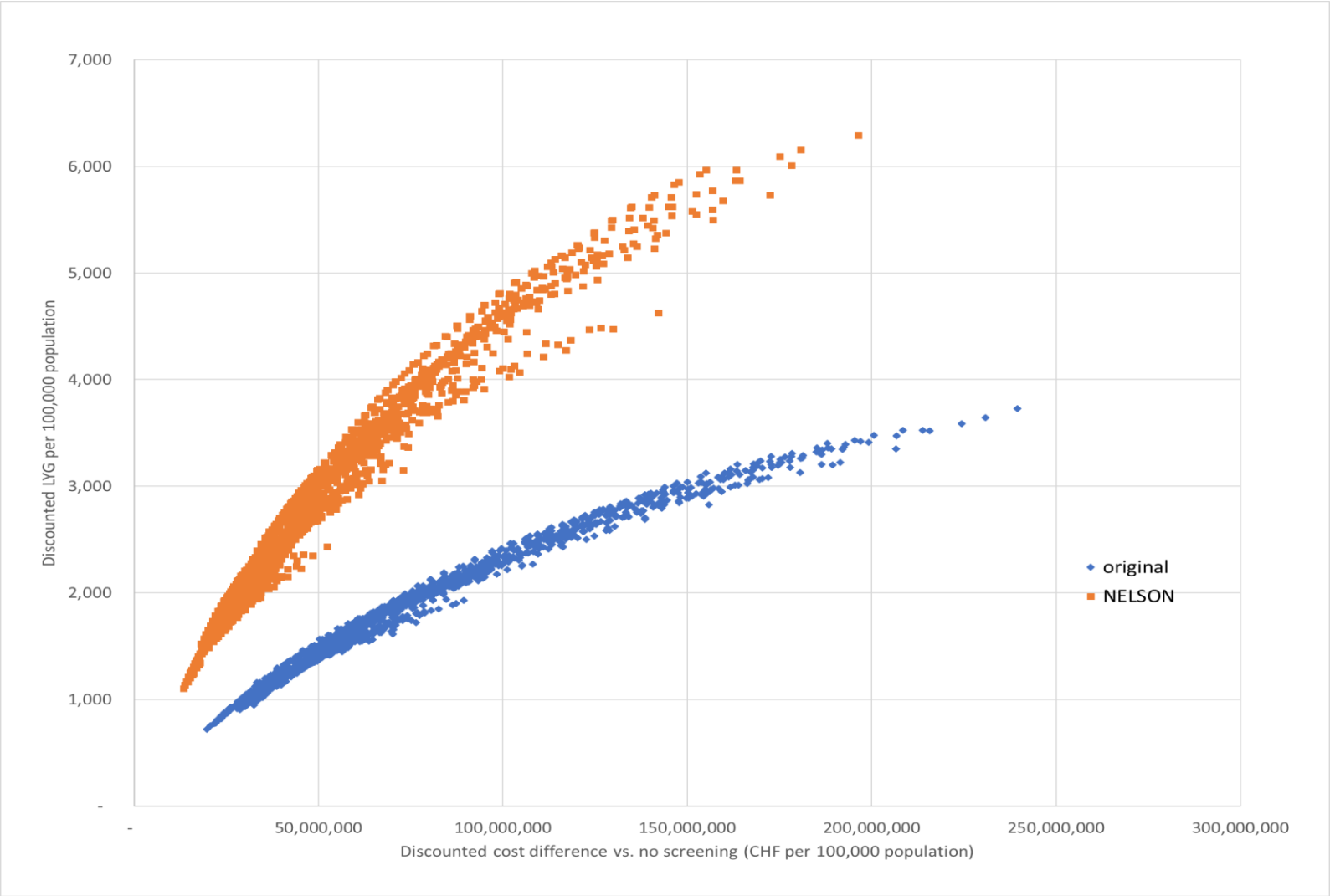


Figure 25 Costs and life-years gained of all screening scenarios versus no screening. Results are presented per 100,000 individuals alive in 2015. Costs and life-years gained were discounted by 3% annually

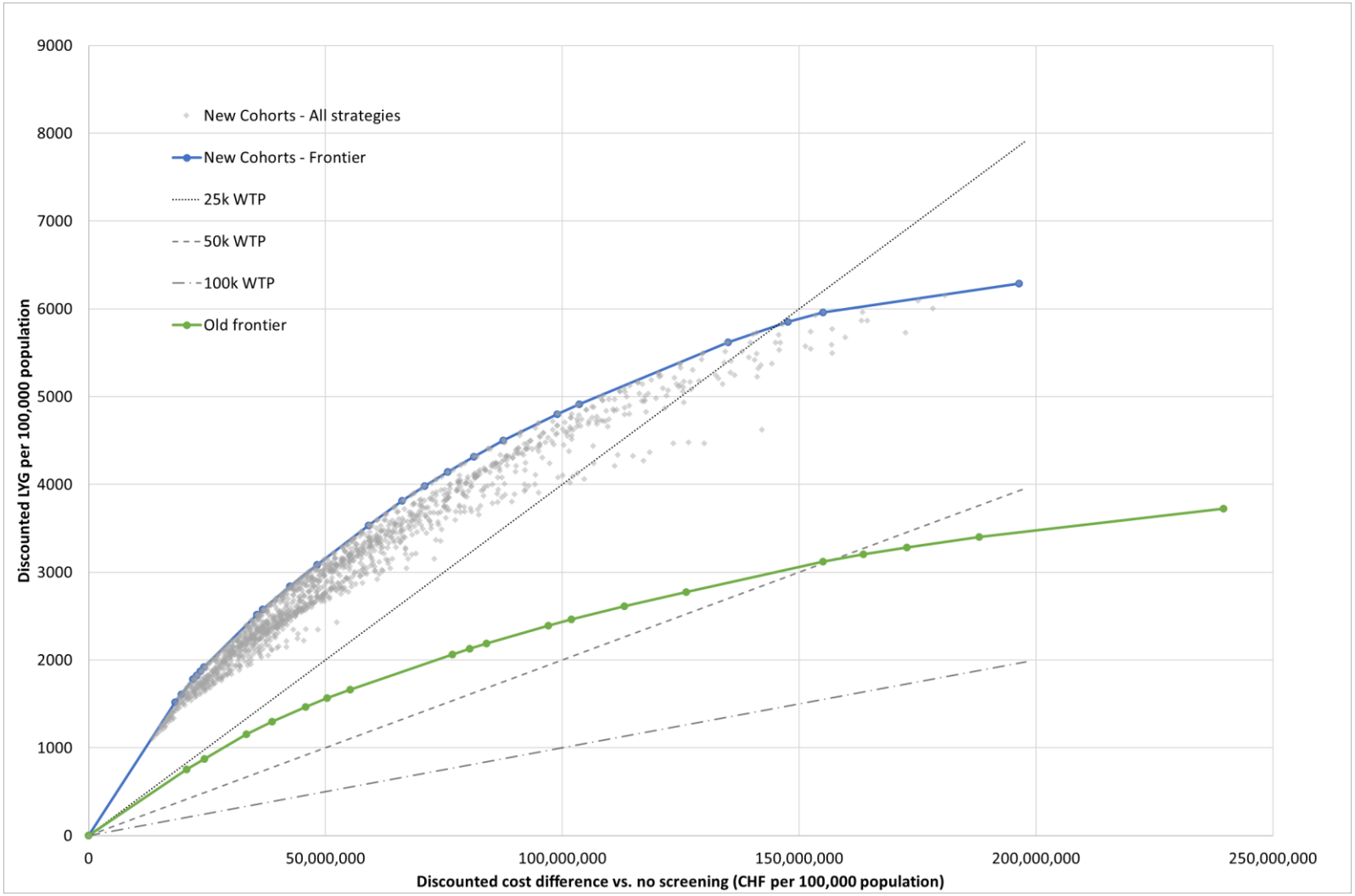


Figure 26 Efficiency frontier for scenarios based on NLST effectiveness (original) and NELSON effectiveness (CT +5%). Results are presented per 100,000 individuals alive in 2015. Costs and life-years gained were discounted by 3% annually

Scenario name	Screening start age	Screening stop age	Smoking criteria / PLCOm2012 risk threshold	Screening interval	Discounted costs compared to no screening per 100,000 (CHF)	Discounted QALYs compared to no screening per 100,000	Discounted costs/QALYs compared to no screening	ICERs (CHF/QALY)
py_60_75_40_10_2	60	75	40 py, max 10 y since cessarion	biennial	18,295,051	1,266	14,452	14,452
plco_60_75_0.027_2	60	75	2.70%	biennial	22,059,363	1,483	14,876	17,349
plco_60_75_0.026_2	60	75	2.60%	biennial	22,777,076	1,520	14,987	19,451
plco_60_75_0.025_2	60	75	2.50%	biennial	23,570,464	1,559	15,114	19,995
plco_60_75_0.024_2	60	75	2.40%	biennial	24,322,828	1,596	15,239	20,552
plco_60_75_0.030_1	60	75	3.00%	annual	35,518,366	2,096	16,949	22,415
plco_60_75_0.029_1	60	75	2.90%	annual	36,749,957	2,145	17,131	24,824
plco_60_75_0.025_1	60	75	2.50%	annual	42,200,738	2,339	18,038	28,052
plco_55_75_0.023_1	55	75	2.30%	annual	45,983,387	2,472	18,603	28,571
plco_60_80_0.025_1	60	80	2.50%	annual	66,162,245	3,137	21,090	30,333
plco_55_80_0.023_1	55	80	2.30%	annual	70,923,118	3,280	21,625	33,393
plco_55_80_0.021_1	55	80	2.10%	annual	75,763,363	3,413	22,197	36,245
plco_55_80_0.019_1	55	80	1.90%	annual	81,314,813	3,562	22,830	37,384

plco_55_80_0.017_1	55	80	1.70%	annual	87,582,791	3,719	23,552	39,912
plco_55_80_0.014_1	55	80	1.40%	annual	98,905,956	3,970	24,910	44,985
plco_50_80_0.013_1	50	80	1.30%	annual	103,593,503	4,066	25,475	48,844
plco_50_80_0.012_1	50	80	1.20%	annual	108,652,241	4,151	26,172	59,543
plco_50_85_0.013_1	50	85	1.30%	annual	135,054,619	4,587	29,446	60,672
plco_50_85_0.011_1	50	85	1.10%	annual	147,669,710	4,781	30,889	65,024
nelson_50_85_25_10_30_5_25_1	50	85	25 y x 10 cig/day or 30 y x 5 cig/day, max 25 y since cessation	annual	175,144,828	5,009	34,965	120,183
py_50_85_10_25_1	50	85	10 py, max 25 since cessation	annual	196,370,769	5,173	37,959	129,444

Results are presented per 100,000 individuals alive in 2021. Costs and life-years gained are discounted by 3% annually.
Explanation of scenario names
Example 1: py_55_75_40_10_2: Inclusion criteria based on NLST (pack-years, PY), start age 55, stop age 75, 40 pack-years, max 10 years since smoking cessation, biennial screening.
Example 2: nelson_50_85_25_10_30_5_10_2: Inclusion criteria based on NELSON, start age 50, stop age 85, 25 years of smoking at least 10 cigarettes per day or 30 years of smoking at least 5 cigarettes per day, max 25 years since smoking cessation, biennial screening.
Example 3: screening_plco_60_75_0.030_1: Inclusion criteria based on risk-assessment (PLCOm2012), start age 60, stop age 75, risk-threshold 3.0%, annual screening.

Table 25 Characteristics and cost-effectiveness of the scenarios based on NELSON effectiveness on the efficiency frontier (including NLST-like, NELEON-like, and risk-based scenarios)

Scenario name	% of the pop. screened	Number of CT screens	Number of CT scans	Number of biopsies	False positive screens	Number of lung cancer cases	Number of lung cancer deaths avoided	Over-diagnosed lung cancers	Number needed to screen per LYG *	Number needed to screen per death avoided *
py_60_75_40_10_2	6.8%	31,536	2,950	146	360	6,799	203	15	2	33
plco_60_75_0.027_2	10.0%	40,172	3,966	191	464	6,802	248	18	3	40
plco_60_75_0.026_2	10.2%	41,876	4,104	198	483	6,803	254	18	3	40
plco_60_75_0.025_2	10.5%	43,722	4,260	206	503	6,803	259	19	3	41
plco_60_75_0.024_2	10.8%	45,551	4,414	214	524	6,803	264	19	3	41
plco_60_75_0.030_1	10.0%	74,586	8,478	365	992	6,813	361	29	2	28
plco_60_75_0.029_1	10.2%	77,625	8,937	382	1,042	6,814	367	29	2	28
plco_60_75_0.025_1	11.2%	91,123	10,986	457	1,262	6,815	396	31	2	28
plco_55_75_0.023_1	11.9%	100,069	12,466	509	1,416	6,816	412	32	2	29
plco_60_80_0.025_1	14.3%	142,965	20,022	759	2,202	6,868	643	84	2	22
plco_55_80_0.023_1	15.0%	154,522	21,928	824	2,410	6,870	664	86	2	23
plco_55_80_0.021_1	15.8%	166,775	23,932	893	2,628	6,872	685	88	2	23
plco_55_80_0.019_1	16.7%	181,046	26,224	972	2,881	6,874	707	89	2	24

plco_55_80_0.017_1	17.6%	197,431	28,823	1,062	3,173	6,875	732	91	2	24
plco_55_80_0.014_1	19.4%	227,024	33,507	1,224	3,701	6,877	772	93	2	25
plco_50_80_0.013_1	20.1%	239,314	35,442	1,291	3,924	6,878	786	94	2	26
plco_50_80_0.012_1	20.8%	252,356	37,500	1,362	4,159	6,880	798	95	2	26
plco_50_85_0.013_1	22.6%	313,729	47,229	1,682	5,331	6,995	1,045	211	2	22
plco_50_85_0.011_1	24.3%	347,788	52,504	1,863	5,955	7,001	1,081	216	2	22
nelson_50_85_25_10_30_5_25_1	21.9%	413,149	64,060	2,195	7,495	6,976	1,044	192	2	21
py_50_85_10_25_1	27.0%	465,115	71,257	2,473	8,290	6,981	1,071	196	2	25

*: number needed to screen refers to the number of individuals screened, rather than the number of screening events.
Explanation of scenario names
Example 1: py_55_75_40_10_2: Inclusion criteria based on NLST (pack-years, PY), start age 55, stop age 75, 40 pack-years, max 10 years since smoking cessation, biennial screening.
Example 2: nelson_50_85_25_10_30_5_10_2: Inclusion criteria based on NELSON, start age 50, stop age 85, 25 years of smoking at least 10 cigarettes per day or 30 years of smoking at least 5 cigarettes per day, max 25 years since smoking cessation, biennial screening.
Example 3: plco_60_75_0.030_1: Inclusion criteria based on risk-assessment (PLCOm2012), start age 60, stop age 75, risk-threshold 3.0%, annual screening.

Table 26 Overview of selected outcomes (per 100,000 individuals alive in 2015) for scenarios on the overall efficiency frontier (not discounted)

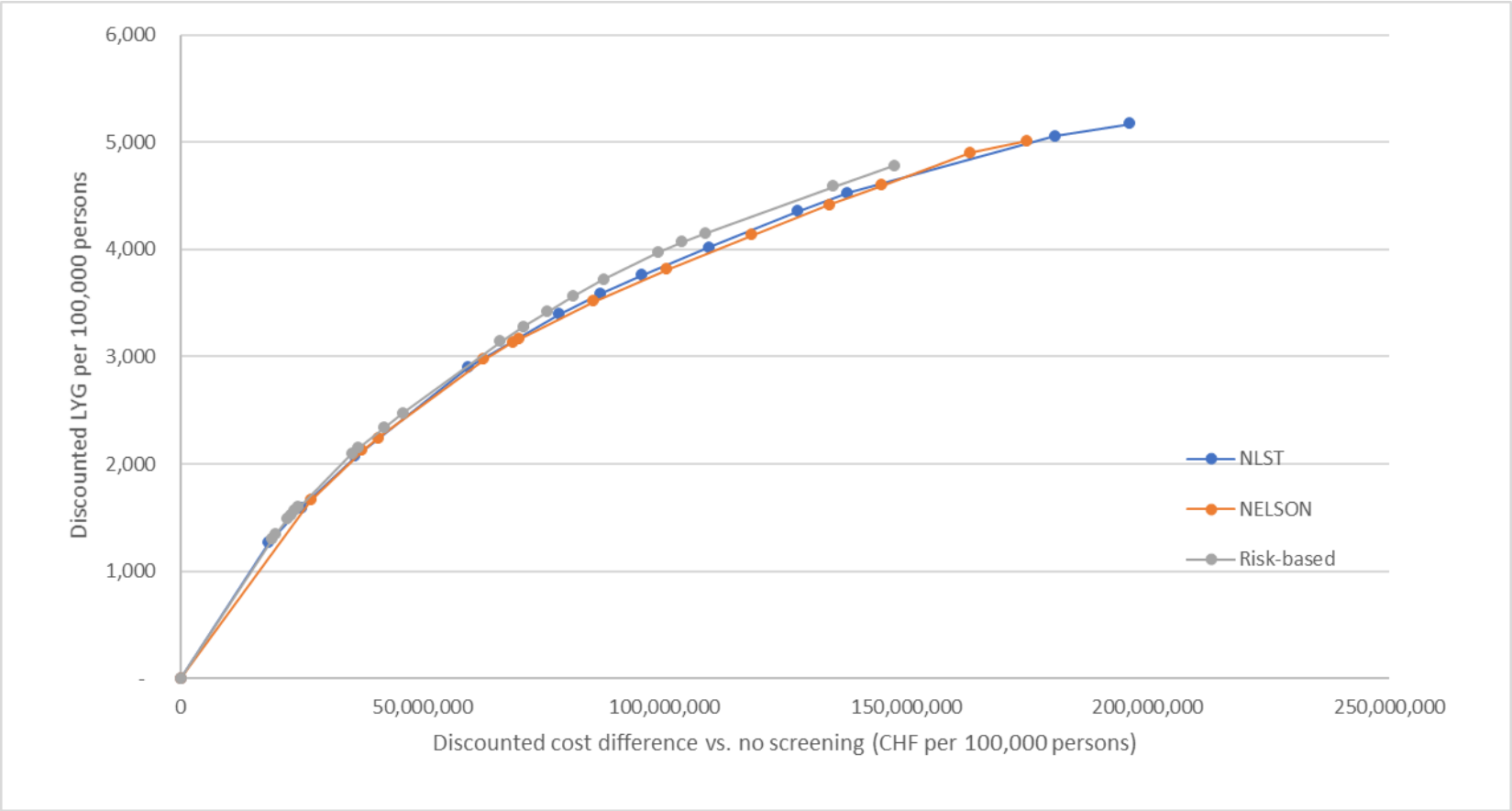


Figure 27 Efficiency frontiers of NLST vs. NELSON vs. risk-based scenarios (based on costs/LYG)

Scenario name	Screening start age	Screening stop age	Smoking criteria	Screening interval	Max years since smoking cessation	Discounted costs compared to no screening per 100,000 (CHF)	Discounted QALYs compared to no screening per 100,000	Discounted costs/QALYs compared to no screening	ICERs (CHF/QALY)
py_60_75_40_10_2	60	75	40 py	biennial	10	18,295,051	1,266	14,452	14,452
py_60_75_30_10_2	60	75	30 py	biennial	10	24,993,150	1,589	15,731	20,745
py_60_80_30_10_2	60	80	30 py	biennial	10	36,060,061	2,071	17,412	22,953
py_60_80_30_10_1	60	80	30py	annual	10	59,495,927	2,905	20,483	28,110
py_60_80_20_10_1	60	80	20 py	annual	10	78,242,225	3,396	23,040	38,156
py_60_80_20_15_1	60	80	20 py	annual	15	86,884,265	3,586	24,231	45,556
py_60_80_20_20_1	60	80	20 py	annual	20	95,355,132	3,757	25,383	49,569
py_55_80_20_20_1	55	80	20 py	annual	20	109,256,040	4,015	27,210	53,748
py_55_85_20_20_1	55	85	20 py	annual	20	127,586,612	4,352	29,315	54,389
py_55_85_20_25_1	55	85	20 py	annual	25	137,970,206	4,526	30,487	59,931
py_55_85_10_25_1	55	85	10 py	annual	25	180,815,108	5,053	35,787	81,299
py_50_85_10_25_1	50	85	10 py	annual	25	196,370,769	5,173	37,959	128,923

*: number needed to screen refers to the number of individuals screened, rather than the number of screening events.
Explanation of scenario name
Example: py_60_75_40_10_2: Inclusion criteria based on NLST (pack-years, PY), start age 60, stop age 75, 40 pack-years, max 10 years since smoking cessation, biennial screening

Table 27 Characteristics and cost-effectiveness of the scenarios based on NELSON effectiveness on the efficiency frontier (NLST-like scenarios)

Scenario name	% of the pop. screened	Number of CT screens	Number of CT scans	Number of biopsies	False positive screens	Number of lung cancer cases	Number of lung cancer deaths avoided	Over-diagnosed lung cancers	Number needed to screen per LYG *	Number needed to screen per death avoided *
py_60_75_40_10_2	6.8%	31,536	2,950	146	360	6,799	203	15	2	33
py_60_75_30_10_2	9.6%	47,633	4,368	219	542	6,802	253	17	3	38
py_60_80_30_10_2	11.0%	65,610	6,486	306	793	6,836	407	52	2	27
py_60_80_30_10_1	11.1%	125,250	18,426	677	2,011	6,854	558	70	2	20
py_60_80_20_10_1	14.7%	173,074	25,850	943	2,814	6,863	646	78	2	23
py_60_80_20_15_1	16.2%	195,689	29,503	1,070	3,207	6,867	683	82	2	24
py_60_80_20_20_1	17.6%	217,815	33,101	1,196	3,593	6,870	717	85	2	25
py_55_80_20_20_1	18.0%	249,707	38,108	1,355	4,247	6,870	739	85	2	24
py_55_85_20_20_1	18.6%	288,152	44,107	1,546	5,024	6,957	915	172	2	20
py_55_85_20_25_1	19.6%	315,432	48,553	1,694	5,542	6,965	959	181	2	20
py_55_85_10_25_1	26.8%	430,726	66,007	2,307	7,571	6,981	1,062	196	2	25
py_50_85_10_25_1	27.0%	465,115	71,257	2,473	8,290	6,981	1,071	196	2	25

*: number needed to screen refers to the number of individuals screened, rather than the number of screening events.
Explanation of scenario name
Example: py_60_75_40_10_2: Inclusion criteria based on NLST (pack-years, PY), start age 60, stop age 75, 40 pack-years, max 10 years since smoking cessation, biennial screening

Table 28 Overview of selected outcomes (per 100,000 individuals alive in 2015) for NLST-like scenarios on the efficiency frontier (not discounted)

Scenario name	Screening start age	Screening stop age	Smoking criteria	Screening interval	Discounted costs compared to no screening per 100,000 (CHF)	Discounted QALYs compared to no screening per 100,000	Discounted costs/QALYs compared to no screening	ICERs (CHF/QALY)
nelson_60_75_30_15_35_10_10_2	60	75	30 y x 15 cig/day or 35 y x 10 cig/day	biennial	26,967,923	1,661	16,236	16,236
nelson_60_80_30_15_35_10_10_2	60	80	30 y x 15 cig/day or 35 y x 10 cig/day	biennial	37,613,171	2,121	17,734	23,142
nelson_60_80_25_15_30_10_10_2	60	80	25 y x 15 cig/day or 30 y x 10 cig/day	biennial	40,877,390	2,240	18,249	27,425
nelson_60_80_30_15_35_10_10_1	60	80	30 y x 15 cig/day or 35 y x 10 cig/day	annual	62,722,725	2,974	21,089	29,753
nelson_60_80_25_15_30_10_10_1	60	80	25 y x 15 cig/day or 30 y x 10 cig/day	annual	68,670,012	3,135	21,903	36,955
nelson_60_80_20_15_25_10_10_1	60	80	20 y x 15 cig/day or 25 y x 10 cig/day	annual	69,899,133	3,164	22,090	42,280
nelson_55_80_30_15_35_10_25_1	55	80	30 y x 15 cig/day or 35 y x 10 cig/day	annual	85,420,057	3,516	24,292	44,067
nelson_55_80_25_15_30_10_25_1	55	80	25 y x 15 cig/day or 30 y x 10 cig/day	annual	100,597,579	3,811	26,400	51,599
nelson_55_85_25_15_30_10_25_1	55	85	25 y x 15 cig/day or 30 y x 10 cig/day	annual	118,150,246	4,132	28,592	54,563
nelson_60_85_25_10_30_5_20_1	60	85	25 y x 10 cig/day or 30 y x 5 cig/day	annual	134,157,792	4,416	30,379	56,389
nelson_60_85_25_10_30_5_25_1	60	85	25 y x 10 cig/day or 30 y x 5 cig/day	annual	145,070,746	4,599	31,544	59,706
nelson_55_85_25_10_30_5_25_1	55	85	25 y x 10 cig/day or 30 y x 5 cig/day	annual	163,352,702	4,899	33,347	61,008

Explanation of scenario codes
Example: lc_screening_nelson_55_80_30_15_35_10_10_2: Inclusion criteria based on NELSON, start age 55, stop age 80, 30 years of smoking at least 15 cigarettes per day or 35 years of smoking at least 10 cigarettes per day, max 10 years since smoking cessation, biennial screening.

Table 29 Characteristics and cost-effectiveness of the scenarios based on NELSON effectiveness on the efficiency frontier (NELSON-like scenarios)

Scenario name	% of the pop. screened	Number of CT screens	Number of CT scans	Number of biopsies	False positive screens	Number of lung cancer cases	Number of lung cancer deaths avoided	Over-diagnosed lung cancers	Number needed to screen per LYG *	Number needed to screen per death avoided *
nelson_60_75_30_15_35_10_10_2	10.0%	53,212	4,769	242	602	6,802	263	18	3	38
nelson_60_80_30_15_35_10_10_2	11.1%	70,400	7,029	329	860	6,835	410	50	2	27
nelson_60_80_25_15_30_10_10_2	12.4%	77,865	7,812	365	953	6,837	431	53	2	29
nelson_60_80_30_15_35_10_10_1	11.2%	134,487	20,261	734	2,201	6,853	562	69	2	20
nelson_60_80_25_15_30_10_10_1	12.6%	148,646	22,353	811	2,430	6,856	591	72	2	21
nelson_60_80_20_15_25_10_10_1	12.9%	151,249	22,664	824	2,465	6,857	596	72	2	22
nelson_55_80_30_15_35_10_25_1	12.6%	189,794	29,609	1,035	3,304	6,861	652	76	1	19
nelson_55_80_25_15_30_10_25_1	15.1%	227,924	35,515	1,244	3,962	6,865	701	81	2	22
nelson_55_85_25_15_30_10_25_1	15.6%	264,384	41,064	1,420	4,699	6,947	866	163	2	18
nelson_60_85_25_10_30_5_20_1	20.9%	314,800	48,523	1,708	5,435	6,966	960	182	2	22
nelson_60_85_25_10_30_5_25_1	21.7%	344,624	53,883	1,879	6,028	6,976	1,010	192	2	21
nelson_55_85_25_10_30_5_25_1	21.9%	387,033	60,446	2,082	6,939	6,976	1,036	192	2	21
nelson_50_85_25_10_30_5_25_1	21.9%	413,149	64,060	2,195	7,495	6,976	1,044	192	2	21

*: number needed to screen refers to the number of individuals screened, rather than the number of screening events.
Explanation of scenario codes
Example: lc_screening_nelson_55_80_30_15_35_10_10_2: Inclusion criteria based on NELSON, start age 55, stop age 80, 30 years of smoking at least 15 cigarettes per day or 35 years of smoking at least 10 cigarettes per day, max 10 years since smoking cessation, biennial screening.

Table 30 Overview of selected outcomes (per 100,000 individuals alive in 2015) for NELSON-like scenarios on the efficiency frontier (not discounted)

Scenario name	Screening start age	Screening stop age	PLCOm2012 risk threshold	Screening interval	Discounted costs compared to no screening per 100,000 (CHF)	Discounted QALYs compared to no screening per 100,000	Discounted costs/QALYs compared to no screening	ICERs (CHF/QALY)
plco_60_75_0.032_2	60	75	3.20%	biennial	18,993,922	1,301	12,111	14,595
plco_60_75_0.031_2	60	75	3.10%	biennial	19,550,543	1,337	12,137	14,618
plco_60_75_0.027_2	60	75	2.70%	biennial	22,059,363	1,483	12,371	14,876
plco_60_75_0.026_2	60	75	2.60%	biennial	22,777,076	1,520	12,466	14,987
plco_60_75_0.025_2	60	75	2.50%	biennial	23,570,464	1,559	12,576	15,114
plco_60_75_0.024_2	60	75	2.40%	biennial	24,322,828	1,596	12,682	15,239
plco_60_75_0.030_1	60	75	3.00%	annual	35,518,366	2,096	14,114	16,949
plco_60_75_0.029_1	60	75	2.90%	annual	36,749,957	2,145	14,270	17,131
plco_60_75_0.025_1	60	75	2.50%	annual	42,200,738	2,339	15,041	18,038
plco_55_75_0.023_1	55	75	2.30%	annual	45,983,387	2,472	15,525	18,603
plco_60_80_0.025_1	60	80	2.50%	annual	66,162,245	3,137	17,352	21,090
plco_55_80_0.023_1	55	80	2.30%	annual	70,923,118	3,280	17,809	21,625
plco_55_80_0.021_1	55	80	2.10%	annual	75,763,363	3,413	18,293	22,197
plco_55_80_0.019_1	55	80	1.90%	annual	81,314,813	3,562	18,833	22,830
plco_55_80_0.017_1	55	80	1.70%	annual	87,582,791	3,719	19,444	23,552
plco_55_80_0.014_1	55	80	1.40%	annual	98,905,956	3,970	20,590	24,910
plco_50_80_0.013_1	50	80	1.30%	annual	103,593,503	4,066	21,070	25,475
plco_50_80_0.012_1	50	80	1.20%	annual	108,652,241	4,151	21,654	26,172
plco_50_85_0.013_1	50	85	1.30%	annual	135,054,619	4,587	24,031	29,446
plco_50_85_0.011_1	50	85	1.10%	annual	147,669,710	4,781	25,231	30,889

Explanation of scenario codes
Example: lc_screening_plco_60_75_0.030_1: Inclusion criteria based on risk-assessment (PLCOm2012), start age 60, stop age 75, risk-threshold 3.0%, annual screening.

Table 31 Characteristics and cost-effectiveness of the scenarios based on NELSON effectiveness on the efficiency frontier (risk-based scenarios)

Scenario name	% of the pop. screened	Number of CT screens	Number of CT scans	Number of biopsies	False positive screens	Number of lung cancer cases	Number of lung cancer deaths avoided	Over-diagnosed lung cancers	Number needed to screen per LYG *	Number needed to screen per death avoided *
plco_60_75_0.032_2	9.0%	32,729	3,355	158	381	6,801	223	17	3	40
plco_60_75_0.031_2	9.2%	34,137	3,474	164	397	6,801	229	17	3	40
plco_60_75_0.027_2	10.0%	40,172	3,966	191	464	6,802	248	18	3	40
plco_60_75_0.026_2	10.2%	41,876	4,104	198	483	6,803	254	18	3	40
plco_60_75_0.025_2	10.5%	43,722	4,260	206	503	6,803	259	19	3	41
plco_60_75_0.024_2	10.8%	45,551	4,414	214	524	6,803	264	19	3	41
plco_60_75_0.030_1	10.0%	74,586	8,478	365	992	6,813	361	29	2	28
plco_60_75_0.029_1	10.2%	77,625	8,937	382	1,042	6,814	367	29	2	28
plco_60_75_0.025_1	11.2%	91,123	10,986	457	1,262	6,815	396	31	2	28
plco_55_75_0.023_1	11.9%	100,069	12,466	509	1,416	6,816	412	32	2	29
plco_60_80_0.025_1	14.3%	142,965	20,022	759	2,202	6,868	643	84	2	22
plco_55_80_0.023_1	15.0%	154,522	21,928	824	2,410	6,870	664	86	2	23
plco_55_80_0.021_1	15.8%	166,775	23,932	893	2,628	6,872	685	88	2	23
plco_55_80_0.019_1	16.7%	181,046	26,224	972	2,881	6,874	707	89	2	24
plco_55_80_0.017_1	17.6%	197,431	28,823	1,062	3,173	6,875	732	91	2	24
plco_55_80_0.014_1	19.4%	227,024	33,507	1,224	3,701	6,877	772	93	2	25
plco_50_80_0.013_1	20.1%	239,314	35,442	1,291	3,924	6,878	786	94	2	26
plco_50_80_0.012_1	20.8%	252,356	37,500	1,362	4,159	6,880	798	95	2	26
plco_50_85_0.013_1	22.6%	313,729	47,229	1,682	5,331	6,995	1,045	211	2	22
plco_50_85_0.011_1	24.3%	347,788	52,504	1,863	5,955	7,001	1,081	216	2	22

*: number needed to screen refers to the number of individuals screened, rather than the number of screening events.
Explanation of scenario codes
Example: lc_screening_plco_60_75_0.030_1: Inclusion criteria based on risk-assessment (PLCOm2012), start age 60, stop age 75, risk-threshold 3.0%, annual screening.Example 3: plco_60_75_0.030_1: Inclusion criteria based on risk-assessment (PLCOm2012), start age 60, stop age 75, risk-threshold 3.0%, annual screening.

Table 32 Overview of selected outcomes (per 100,000 individuals alive in 2015) for risk-based scenarios on the efficiency frontier (not discounted)

17.2.6. Distributions of all scenarios by inclusion criteria (NELSON vs. NLST. vs risk-based)

The distributions of all the scenarios according to inclusion strategy (NLST-like, NELSON-like, risk-based) is illustrated in Figure 28. It can be noticed that the distributions are similar across different inclusion strategies. Strategies based on the PLCOm2012 risk-asessment seem to be slighlt more effective (lead to more QALY gained for lower costs).

17.2.7. Impact of screening intensity (annual, biennial, triennial screening)

The distributions of the scenarios according to screening interval (annual, biennial, triennial) is illustrated in Figure 29. Triennial screening scenarios seems to be less expensive but also less effective if compared to biennial and annual screening strategies. As confirmed from the calculation of the efficiency frontiers, the most cost-effective triennial screenings are dominated from biennial ones.

17.2.8. Impact of smoking intensity

The distributions of the scenarios according to smoking intensity (according to pack-years for the NLST-like scenarios) is illustrated in Figure 30. In general, scenarios for high smoking intensity (i.e. 30-40 pack-years) tend to be closer to the efficiency frontier if compared to those with lower smoking intensity.

17.2.9. Impact of risk assessment

The distributions of the scenarios according to risk-threshold (based on PLCOm2012 risk assessment) is illustrated in Figure 31. It can be noticed that scenarios using a high-risk assessment threshold (> 2.0%) tend to be concentrated in the lower-left part, while scenarios with low-risk threshold are more equally distributed.

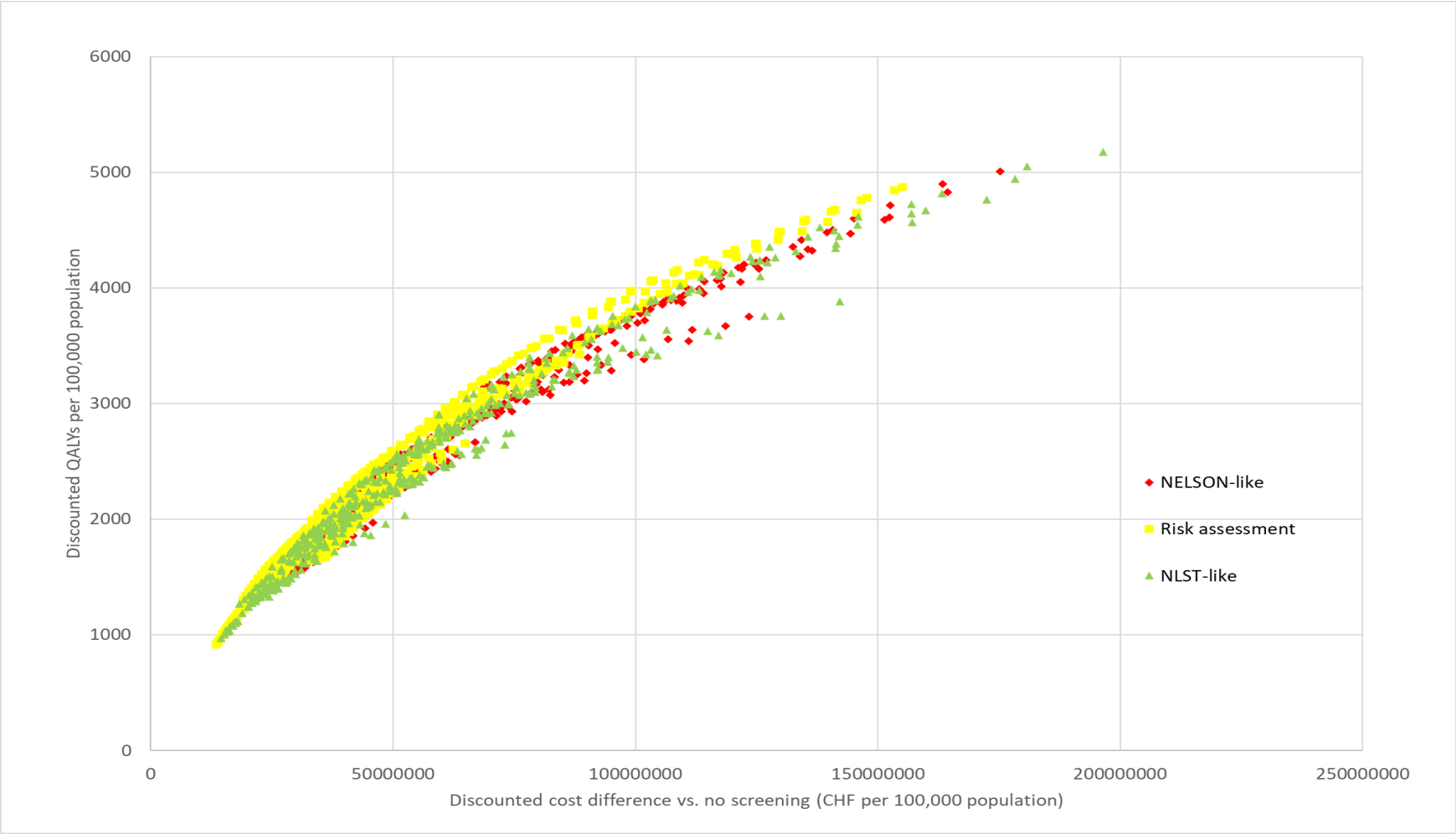


Figure 28 Scenario stratified by selection criteria (NLST vs. NELSON vs. risk-based)

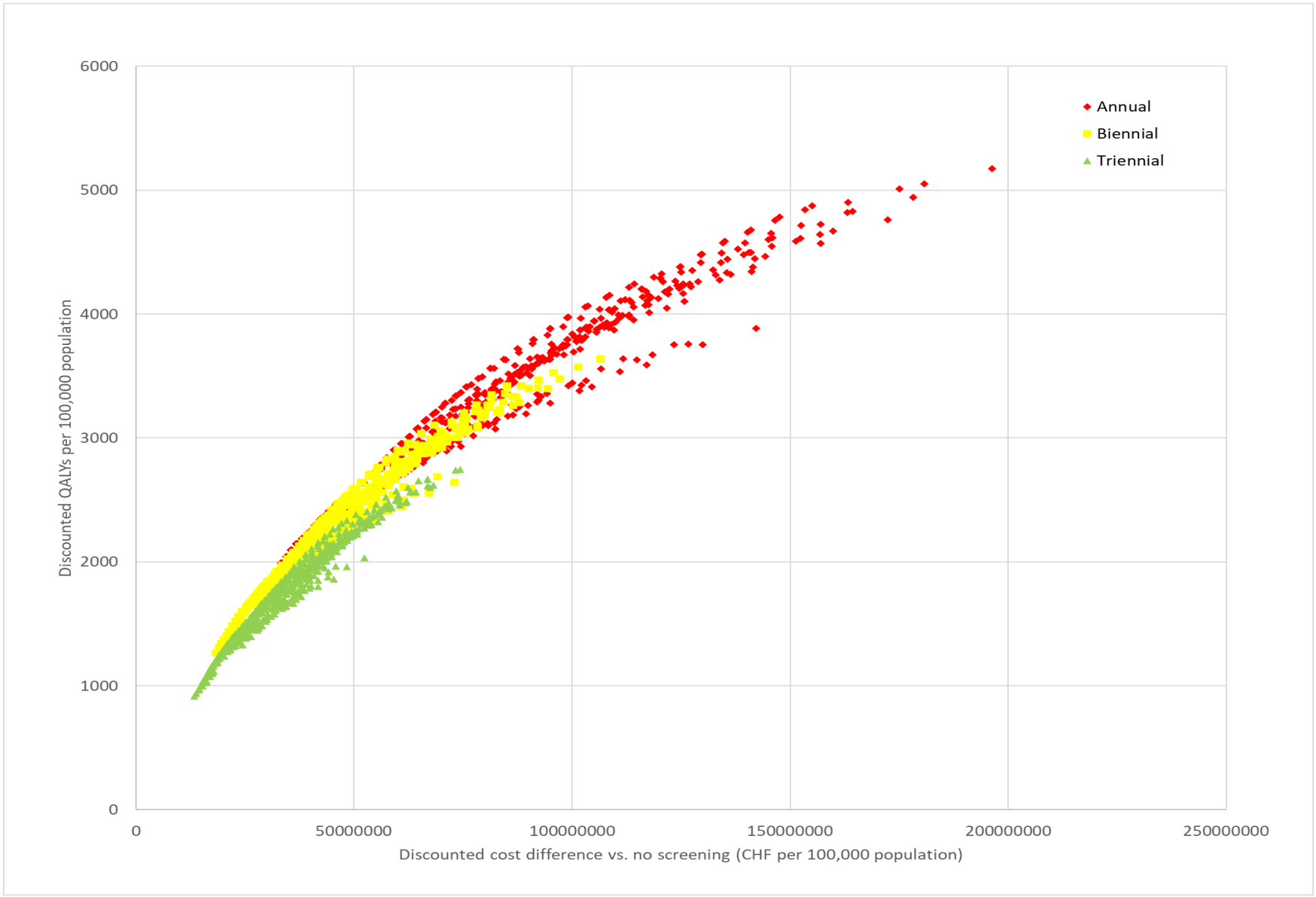


Figure 29 Scenarios by screening interval (annual, biennial, triennial)

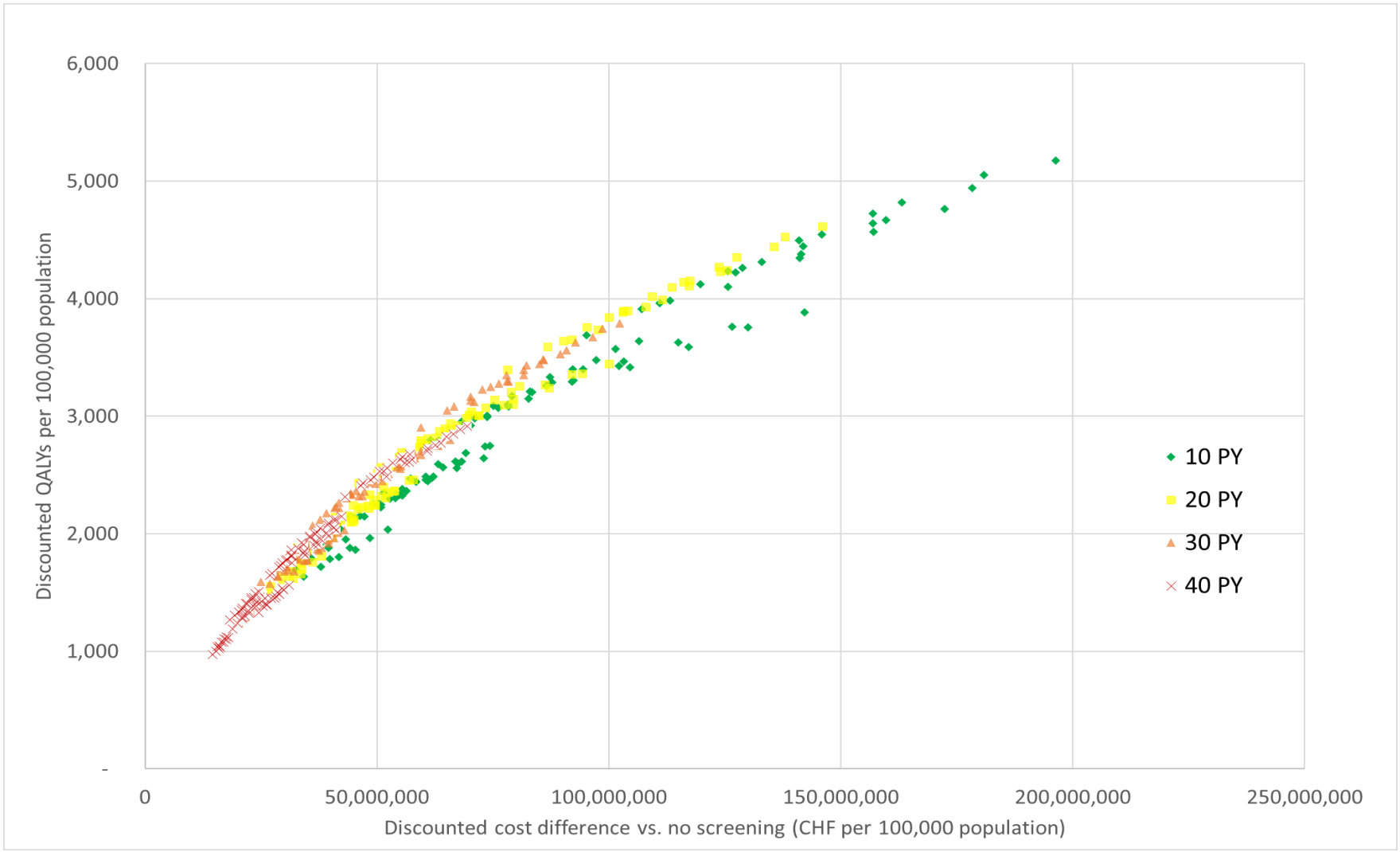


Figure 30 Scenarios based on NELSON effectiveness, stratified by pack-years (only NLST-like scenarios)

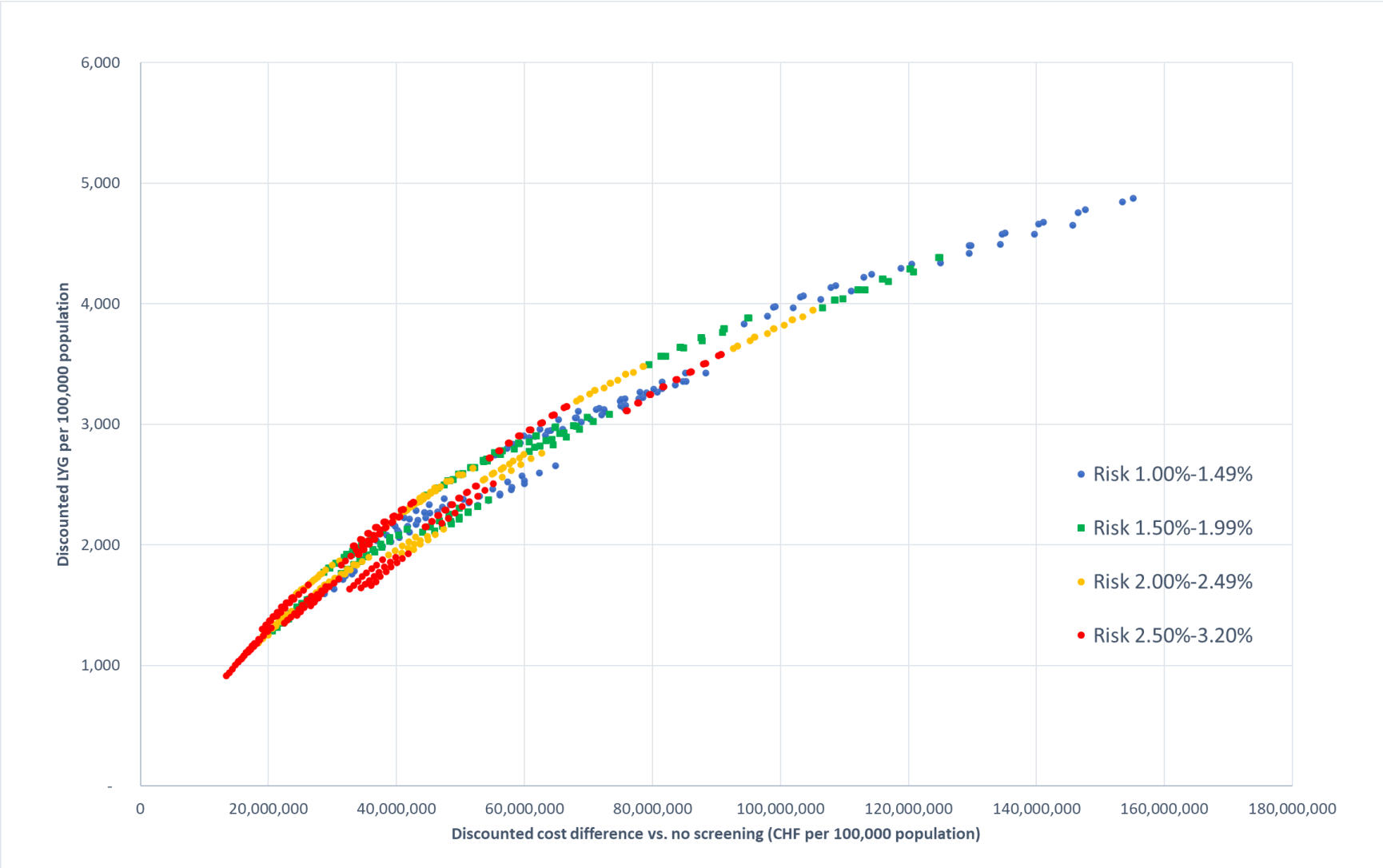


Figure 31 Scenarios based on risk assessment, stratified by risk threshold

17.2.10. Sensitivity analyses

17.2.11. Variation of the LDCT and treatment costs

In the first sensitivity analyses we varied the costs of LDCT and treatment by 30% for all scenarios on the efficiency frontier. As illustrated in Figure 32, all efficiency frontiers with lower costs assumptions generally led to lower ICERs (i.e., to an increased cost-effectiveness), while efficiency frontiers assuming higher costs had an opposite impact. For example, the ICERs of the first scenario on the efficiency frontier ranged between CHF 9,468 per QALY gained if the costs (of both LDCT and treatment) were reduced by 30% and CHF 17,015 per QALY gained if the costs were increased by 30% (the ICER in the main analysis was CHF 14,452 per QALY gained). In general, a variation of all costs by 30% led to a decrease/increase of the ICERs of 33%. As also visible in Figure 32, LDCT screening costs variation had a bigger influence on the ICERs if compared to the treatment cost variation (±25.5% change in the ICERs compared to the base case vs. ±7.2%). The main reason is that a change of LDCT costs has no impact on the costs of the no screening scenario, while a change in the treatment costs results in a cost change for all scenarios. An increase in treatment costs can also partially benefit the cost effectiveness of a screening scenario since it increases the gross cost savings from reduced terminal care for lung cancer. This offsets the increase in costs from initial care and continuing care for lung cancer in a screening scenario.

17.2.12. Additional patient selection considering life expectancy

In this sensitivity analysis an additional screening selection criterion was added: In all modelled scenarios (based on NLST-criteria, NELSON-criteria, or risk assessment) only patients with an estimated life expectancy of at least 5 years were considered eligible for screening (and especially treatment). Figure 33 illustrates the distribution of all scenarios, the new efficiency frontier as well as the frontier calculated in the main analysis. It can be noticed that the efficiency frontier shifts to the left, suggesting that excluding patients with a life expectancy below five years would increase the cost-effectiveness of lung cancer screening. This is consistent with analyses suggesting that incorporating life-expectancy would greatly reduce overdiagnosis, while retaining the life-years gained by screening. [182] The new analyses incorporating life expectancy as additional selection criterion also led to a change of the scenarios on the efficiency frontier (Table 33). If compared with the main analyses, it can be noticed that the new efficiency frontier tends to include scenarios with higher screening stop ages (there are considerably more scenarios with screening stop age at 80 or 85 years, especially in the second half of the frontier). Main reason is the fact that with the exclusion of patients with low life expectancy (i.e., <5 years) it is possible to select the patients that would benefit the most from screening (and, if necessary, from lung cancer treatment). The benefit of selecting an older screening population is no longer offset by the risk of overdiagnosis, yielding higher upper age limits on the efficient frontier.

17.2.13. Variation of the discount rate

Figure 34 illustrates the effects of the variation of the discount rate on the scenarios on the efficiency frontier. Compared with the main analysis based on a 3% discount rate, using a lower discount rate led to a steeper efficiency frontier (i.e., the cost-effectiveness increased). With a 0% discount rate, the ICERs ranged between CHF 8,335 and CHF 67,730 per QALY gained, while for the base-case analysis the ICERs ranged between CHF 14,452 and CHF 129,444 per QALY gained. On the opposite side, scenarios with higher discount rates led to a decrease in the cost-effectivess (i.e. higher costs per QALY gained). For example, with a 6% discount rate, the ICERs ranged between CHF 23,152 and CHF 213,148 per QALY gained.

It is worth mentioning that the impact of discounting was particularly relevant for the estimated QALY gained, while the changes in the estimated costs were less relevant. For example, the estimated number of QALY gained in the first scenario on the efficiency frontier changed from 1,266 per 100,000 persons in the main analysis (discount rate 3%) to 2,493 QALY gained per 100,000 (+97%) using a 0% discount rate and 701 QALY gained per 100,000 using a 6% discount rate (-45%). For the same scenario, the costs changed from CHF 18.3 million per 100,000 in the main analysis to CHF 20.8 million using a 0% discount rate (+14%) and CHF 16.2 million using a 6% discount rate (-11%). Such different impacts of discounting on the costs and QALYs are not surprising and can be explained with the fact that most costs of screening were assumed to happen in the first few years after screening initiation, while the effects on life expectancy and QALYs become evident only on the long-term (see section 12. Budget impact analysis).

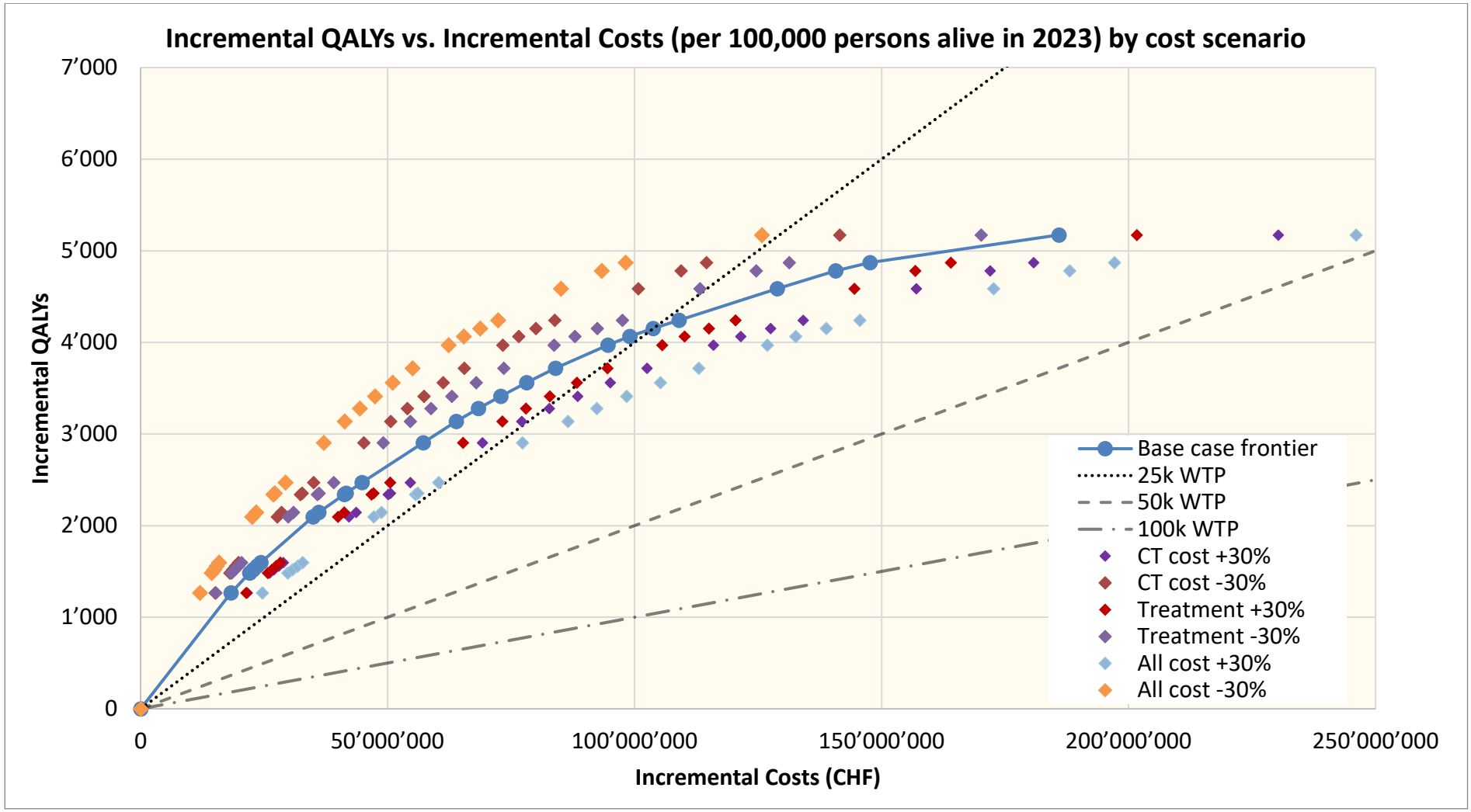


Figure 32 Change in incremental costs and incremental QALYs for scenarios on the efficiency frontier after variating the CT and/or treatment costs by 30%

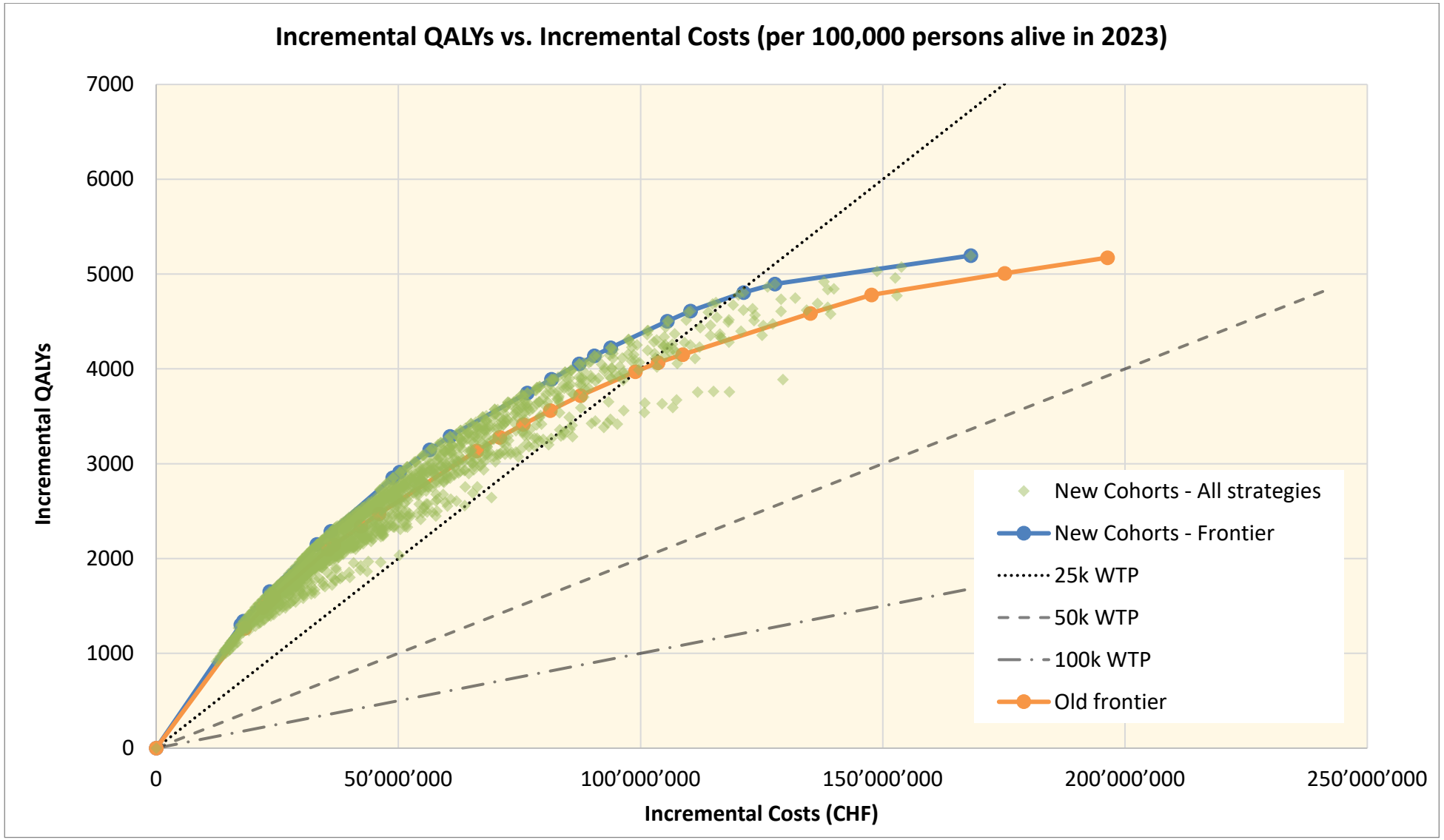


Figure 33 Scenario distribution and efficiency frontier after introduction of life expectancy limit (i.e., only patients with a life expectancy above five years are considered eligible for screening)

Scenario name	Screening start age	Screening stop age	Smoking criteria / PLCOm2012 risk threshold	Screening interval	Discounted costs compared to no screening per 100,000 (CHF)	Discounted QALYs compared to no screening per 100,000	Discounted costs/QALYs compared to no screening	ICERs (CHF/QALY)
py_60_75_40_10_2	60	75	40 py, max 10 y since smoking cessation	biennial	16,977,131	1,268	13,391	13,391
plco_60_75_0.032_2	60	75	3.20%	biennial	17,476,919	1,304	13,406	13,951
plco_60_75_0.031_2	60	75	3.10%	biennial	18,013,596	1,340	13,446	14,902
plco_60_75_0.027_2	60	75	2.70%	biennial	20,408,020	1,485	13,742	16,459
py_60_80_40_10_2	60	80	40 py, max 10 y since smoking cessation	biennial	23,448,474	1,653	14,184	18,092
plco_60_75_0.030_1	60	75	3.00%	annual	32,080,302	2,099	15,280	19,341
plco_60_75_0.024_2	60	75	2.40%	biennial	36,050,510	2,288	15,758	21,094
py_60_80_40_10_1	60	80	40 py, max 10 y since smoking cessation	annual	36,780,987	2,320	15,851	22,310
plco_60_75_0.030_1	60	75	3.00%	annual	48,787,521	2,852	17,105	22,577
plco_60_75_0.029_1	60	75	2.90%	annual	50,220,084	2,913	17,240	23,595
plco_60_75_0.025_1	60	75	2.50%	annual	56,448,930	3,147	17,935	26,558
plco_55_80_0.023_1	55	80	2.30%	annual	60,662,723	3,290	18,436	29,490
plco_55_80_0.021_1	55	80	2.10%	annual	64,926,777	3,424	18,962	31,885
plco_55_80_0.019_1	55	80	1.90%	annual	69,837,827	3,573	19,546	32,968
plco_55_80_0.017_1	55	80	1.70%	annual	75,448,509	3,730	20,228	35,772
plco_55_85_0.019_1	55	85	1.90%	annual	87,286,877	4,055	21,526	36,428
plco_55_85_0.018_1	55	85	1.80%	annual	90,448,954	4,139	21,852	37,536
plco_55_85_0.017_1	55	85	1.70%	annual	93,804,361	4,225	22,204	39,267

plco_55_85_0.014_1	50	85	1.40%	annual	105,497,611	4,504	23,424	41,862
plco_50_85_0.013_1	50	85	1.30%	annual	110,277,548	4,611	23,918	44,789
plco_50_85_0.011_1	50	85	1.10%	annual	121,243,145	4,805	25,232	56,364
plco_50_85_0.01_1	50	85	1.00%	annual	127,712,216	4,896	26,083	70,900
py_50_85_10_25_1	50	85	10 py, max 25 since smoking cessation	annual	168,139,947	5,195	32,363	135,201

Table 33 Scenarios on the efficiency frontier including expected life expectancy as eligibility criterion

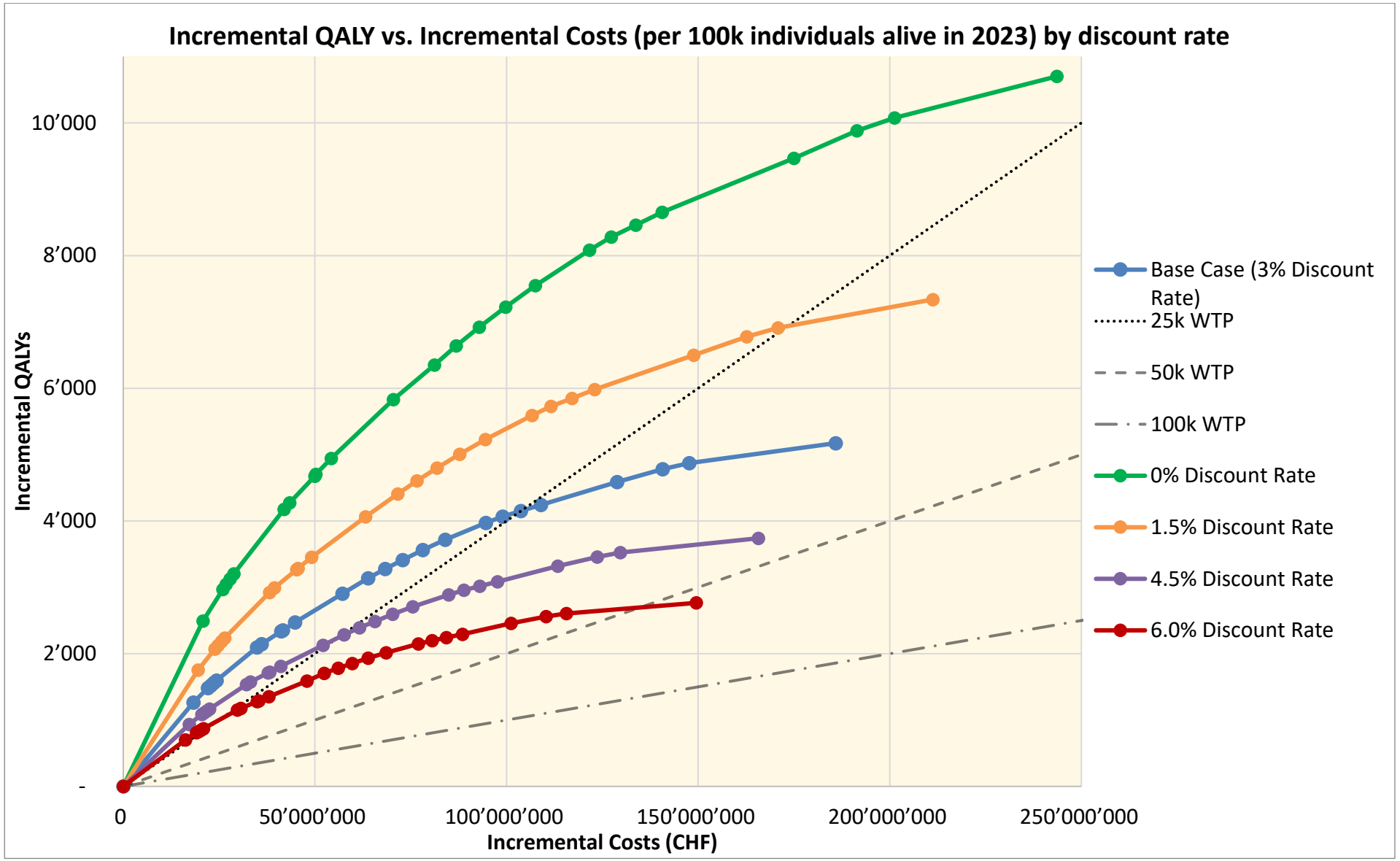


Figure 34 Discount rate variation of the scenarios on the efficiency frontier

17.3. Discussion

The results of the cost-effectiveness analysis suggest that the majority of the investigated lung cancer screening strategies may be cost-effective in Switzerland (assuming a threshold of CHF 100,000 per QALY gained). If compared with no screening, all screening scenarios had a ACER below CHF 62,000 per QALY gained. If only scenario on the efficiency frontier were considered, the ACERs ranged between CHF 14,452 and CHF 37,959 per QALY gained. When compared between each other, only the last two scenarios on the efficiency frontier had an ICER above CHF 100,000 per QALY gained. These scenarios included the largest number of persons (e.g. persons aged between 50 and 85 years with relatively low smoking intensity).

Different smoking eligibility criteria (NLST-like, NELSON-like, or risk-based) for selecting the population to be screened showed comparable results. Nevertheless, patients selection according to PLCOm2012 risk assessment seemed to be slightly more cost-effective than the selection through NLST or NELSON crieria. This was also confirmed by the scenarios on the efficiency frontier, which were predominanty based on the risk-assessment.

The cost-effectiveness was highly dependent on screening intervals and smoking eligibility criteria: Lung cancer screening was generally less costly when performed biannualy (instead of annualy) and when restricted to high-risk individuals. It was particularly cost-effective for smokers or ex-smokers that smoked at least 30/40 pack-years with no more than 10 years since smoking cessation (NLST-like criteria), for smokers or ex-smokers who smoked more than 35/30 years at least 15/10 cigarettes per day, with no more than 10 years since smoking cessation (NELSON-like criteria), and for persons above a PLCOm2012 risk threshold of 1.4%.

Although being more expensive than biennial and triennial screening strategies, annual screening showed the greatest potential reduction in lung cancer mortality and the highest increase of QALYs. For example, the number of lung cancer deaths prevented with screening ranged between 203 per 100,000 persons for the first scenario on the efficiency frontier (biennial screening) to 1,071 per 100,000 persons for the last scenario on the efficiency frontier (annual screening). Therefore, from a public health perspective it may be worth to consider a strategy ‘further up’ on the efficiency frontier to reach higher benefits.

The fact that the scenarios on the efficiency frontier mostly included high intensity smokers (or persons with a high risk based on the PLCOm2012 risk assessment) emphasizes the importance of screening eligibility criteria and risk assessment.

The costs of screening for lung cancer relative to no screening was estimated to be particularly sensitive to changes in the costs of the CT screen, and less to the costs of lung cancer treatment. At any considered cost level, we still find some strategies on the frontier with ICERs below CHF 40,000 per QALY gained. Lung cancer

screening shows improved cost-effectiveness, as well as greatly reduced overdiagnosis, in a scenario in which individuals with a low remaining life expectancy (<5 years) are excluded from screening. It should be noted that this is an idealised scenario of perfect information, but that self-selection into screening by individuals with existing morbidities can improve the cost-effectiveness of organised lung-cancer screening in the direction shown by this sensitivity analysis.

If compared to our previous cost-effectiveness analysis based on NLST effectiveness, the results of the present study based on NELSON suggest that lung cancer screening may be even more cost-effective. Several factors may have contributed to this results. First, LDCT sensitivity in NELSON was higher than in NLST. Based on the LDCT sensitivity differences between NLST and NELSON, in the model we increased the LDCT sensitivity by 5 percentage points for stage IA, IB, and II. This change increased the probability of being diagnosed at an early stage, leading to lower treatment costs and longer life expectancy. Second, the false-positive rate in NELSON was much lower than in NLST. As consequence, the number of unnecessary follow-up tests (e.g. CT examination or biopsies) was strongly reduced. Third, in this cost-effectiveness model we newly implemented the costs for immunotherapy. Immunatherapy is a very expensive treatment that is mainly provided in advanced lung cancer stages (mainly stage IV / terminal care phase). Considering that lung cancer screening lead to a stage shift (i.e. patients undergoing screening are more frequently diagnosed at early stages, and have lower probabilities to progress to advanced/terminal stages), terminal care costs were considerably higher in the no-screening scenario.

One major strength of this cost-effectiveness analysis is the use of effectiveness assumptions based on the NELSON study conducted in Europe. As already mentioned, the NELSON study confirmed the results of the NLST trial concluding that LDCT screening reduces lung cancer mortality. Major differences between NELSON and NLST were higher LDCT sensitivity and lower false-positive rates. Another major strength is the use of new Swiss demographic, epidemiologic, and smoking behaviour data. Compared to our previous work, our updated model included a larger cohort (1940-1979 instead of 1935-1969). Lung cancer epidemiological data (i.e., incidence, mortality, survival), previously based on a 10-year period (2004-2013) included patients diagnosed between 2004 and 2018. A larger sample presumably led to more robust estimations. Concerning smoking behaviour, the data we used in the analyses we published in 2018 were based on five Swiss Health Surveys (1992, 1997, 2002, 2007, and 2012). In the present work we additionally had the data collected in 2017. Another relevant update concerns the estimations of the smoking intensity. In our previous study we only considered cigarette consumptions, while in the present analyses we also included cigars, cigarillos, and pipes (assuming that 1 cigar is equivalent to 2 cigarettes, 1 pipe is equivalent to 2.5 cigarettes, and 1 cigarillos is equivalent to 1 cigarette). Another important update concerns the costs. In our previous cost-effectiveness analysis, costs for treatment (initial care, continuing care, and terminal care), for LDCT screening or follow-up examination, and for biopsy were included. Costs for treatment were based on patient diagnosed and

treated at the University Hospital Zurich between 2011 and 2015. At that time immunotherapy for lung cancer was not available. In the present analyses we newly implemented immunotherapy costs as part of the terminal care costs (assuming that most patients receiving it are in advanced lung cancer stages). In addition, we also included costs related to screening invitation and risk-assessment for screening eligibility. A final strength of this study is the implementation of utility values for the Swiss general population and for different lung cancer stages in order to estimate QALYs (instead of LYG).

This study has several limitations. First, the perspective of cost assessment did not include indirect costs (i.e., costs related to work loss and reduced productivity). A considerable proportion of lung cancer cases occurs in the working population (most scenarios on the efficiency frontier had a screening starting age of 55 or 60). Early detection through screening with a higher probability of remission and cure may lead to lower indirect costs, i.e. additional benefit at the societal level.

Second, accurate information on immunotherapy frequency and costs were not available. According to experts from the University Hospital Zurich, 80% of the patients with metastatic non-small cell lung cancer and 98% of those with metastatic small-cell lung cancer receive immunotherapy. The costs were estimated to be approximatively CHF 5,000 every three weeks. In our model we assumed that immunotherapy is most likely part of the terminal care phase. Therefore, we increased the terminal care phase costs we used in our previous study accordingly. Terminal care in our model was limited to a maximum of 6 months. However, immunotherapy may be provided for longer periods (up to two years). This means that we may have underestimated the real costs of immunotherapy.

18.Budget impact analysis

The aim of the budget impact analysis was to investigate the economic impact of different LDCT screening programs for lung cancer in high-risk populations in Switzerland. The analysis was based on the results of the cost-effectiveness analysis.

18.1. Methods

The budget impact analysis (BIA) was based on the results of the cost-effectiveness analysis.

For each scenario, undiscounted costs outcomes (i.e., costs for LDCT screen and invitation, LDCT follow-up, risk-assessment, biopsies, initial care phase, continuous care phase, and terminal care phases) were calculated by calendar year.

It was assumed that screening would start in 2023, and yearly costs were estimated until 2037 (a relatively long-time horizon was necessary to capture/illustrate the potential benefit of lung cancer on terminal care costs). The modelled population included all persons born between 1940 and 1979 who were estimated to be alive in 2023 (i.e., 4,079,544 persons).

The budget impact of all scenarios was compared to no screening. However, in this BIA we focused a selection of the most cost-effective scenarios (i.e., scenario on the efficiency frontiers). It was assumed that most of the eligible persons would undergo risk assessment in the first screening years. After all individuals in the cohorts of interest have reached the minimum age of eligibility, no further risk assessment costs are assumed.

As in the cost-effectiveness analysis, a healthcare payer perspective was adopted (meaning that any type of indirect costs, for example those related to productivity losses, were not included).

18.2. Results

18.2.1. No screening scenario

According to the no screening scenario modelled in our cost-effectiveness analysis, the annual costs related to lung cancer treatment of the 1940-1979 cohort in Switzerland between 2023 and 2037 were estimated to increase from CHF 474 millions to CHF 724 millions. This increase in costs is due to the fact that the expected number of persons born between 1940 and 1979 that will potentially develop lung cancer is expected to increase in the next decade. As illustrated in Figure 35, the main cost drivers are the terminal care costs, which represent almost two thirds of the total costs and include the estimated costs of immunotherapy treatment for late-stage lung cancer.

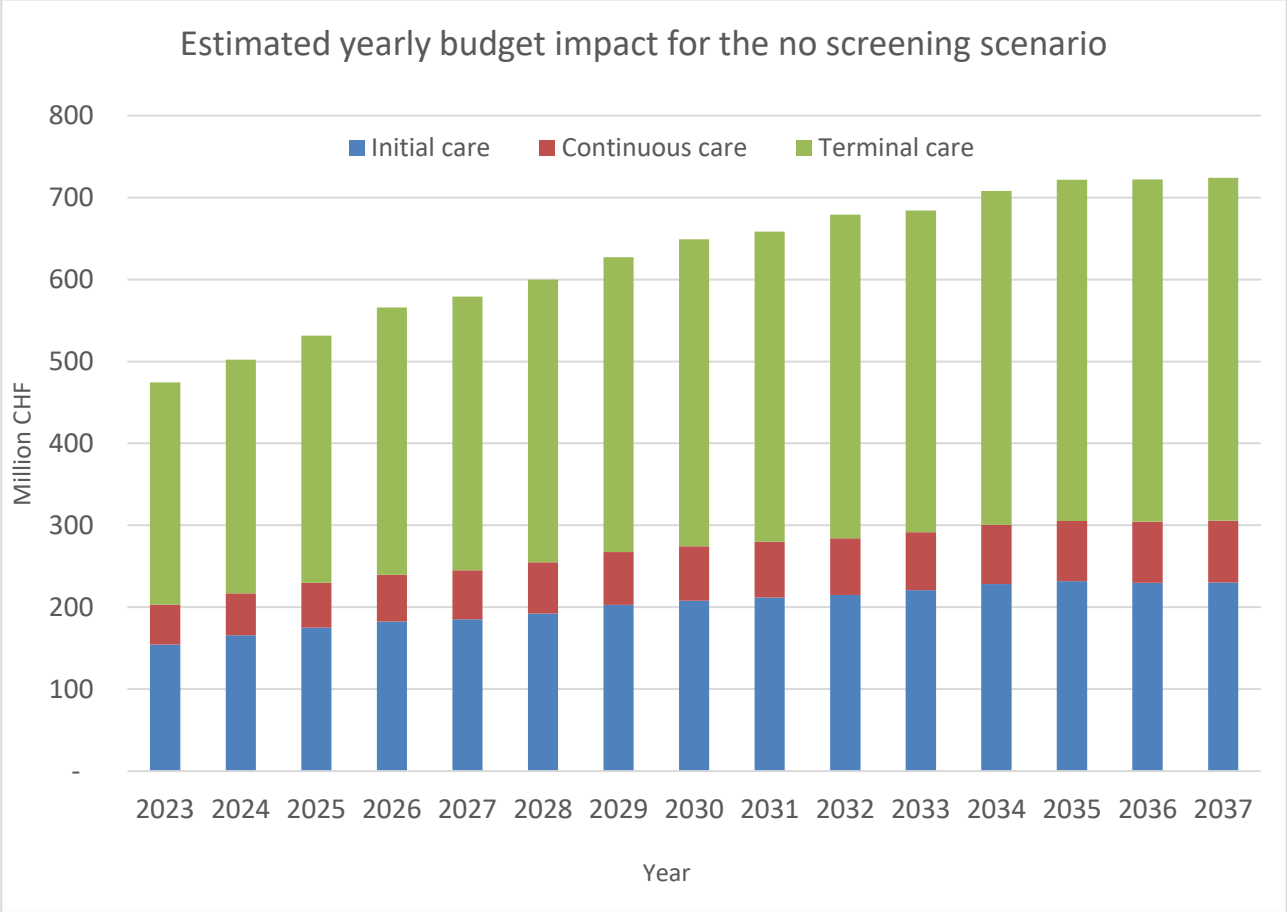


Figure 35 Estimated yearly costs for the no screening scenario between 2023 and 2037

18.2.2. No screening scenario vs. selected screening scenarios

To illustrate the potential budget impact of LDCT screening we compared the no screening scenario with three scenarios on the efficiency frontiers: One based on NLST selection criteria, one based on NELSON selection criteria, and one based on PLCOm2012 risk-assessment. The following screening scenarios were used:

- Scenario based on NLST criteria: Screening age range 60-75 years, 40 pack-years, max 10 years since smoking cessation, biennial screening (scenario name: py_60_75_40_10_2)
- Scenario based on NELSON criteria: Screening age range 60-80 years, at least 15 cig/day for 30 years or 10 cig/day for 35 years, max 10 years since smoking cessation, annual screening (scenario name: nelson_60_80_30_15_35_10_10_1)
- Scenario based on PLCO criteria: Screening range 55-80 years, PLCOm2012 risk threshold 1.7%, annual screening (scenario name plco_55_80_0.017_1)

Figure 36 illustrates the yearly costs of the four scenarios from 2023 to 2037. As mentioned in the previous chapter, we estimated that the cost of the no screening scenario will almost linearly increase until 2037. Compared to no screening, the costs of all screening scenarios are much higher in the first 2-3 years. This is due to the fact that we assumed that risk-assessment and first LDCT screening would be performed in the first 1-2 years. This will automatically lead to the detection and treatment of a high number of lung cancer patients, and thus to considerably higher initial treatment costs.

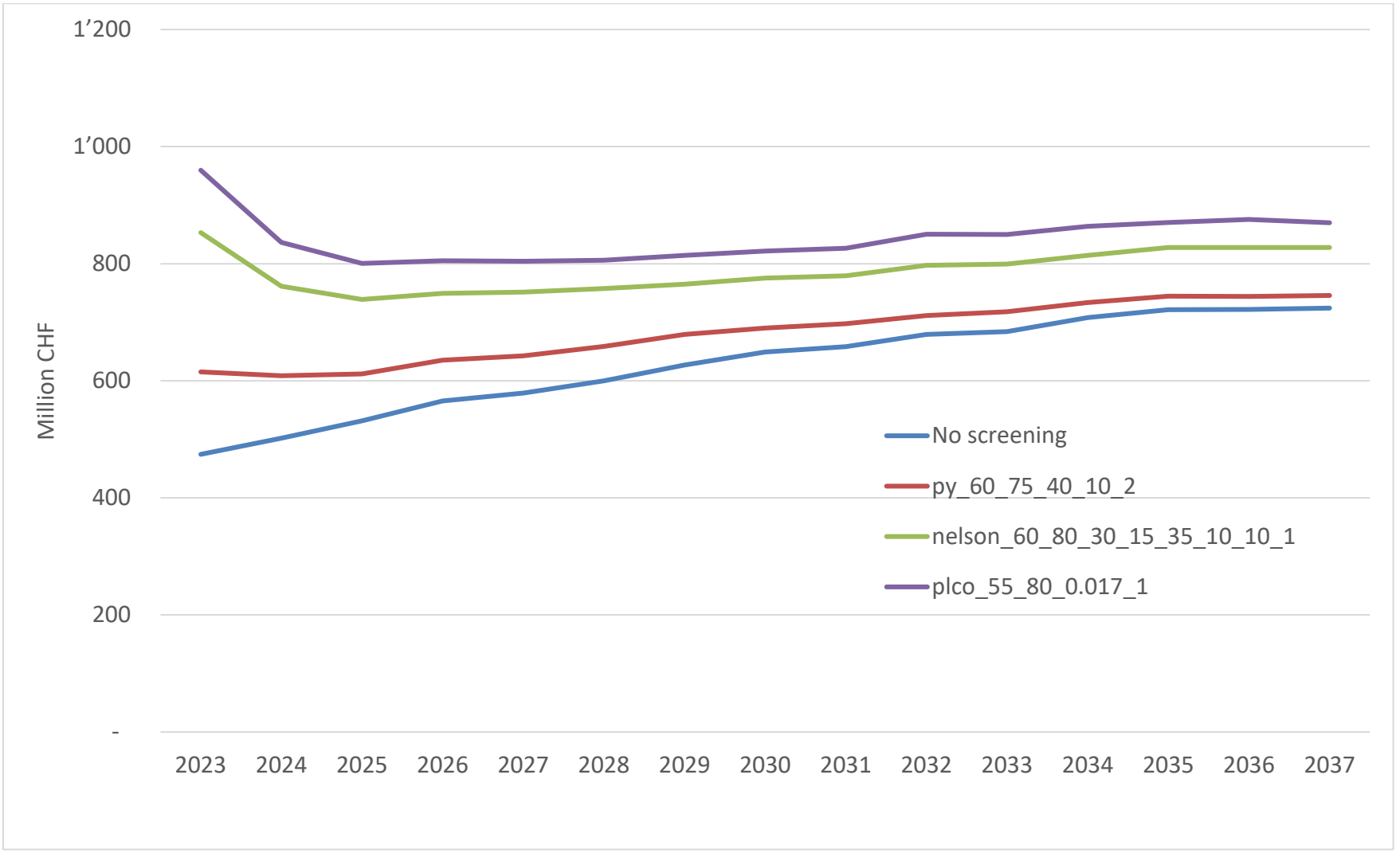


Figure 36 Budget impact of no screening and three screening scenarios between 2023 and 2037

18.2.3. Scenario based on NLST criteria: Screening age range 60-75 years, 40 pack-years, max 10 years since smoking cessation, biennial screening

Figure 37 and Table 34 show the budget impact of the first scenario on the efficiency frontier compared to no screening. As reported in the cost-effectiveness section, in this scenario 6.8% of the population born between 1940 and 1979 would be considered eligible for screening at one point in their lifetime. This means that approximatively 276,000 persons would be screened.

Compared to no screening, this scenario led to higher costs especially in the first years after screening begin. This was mainly due to two strongly related factors. First, most of the costs related to risk assessment were allocated in the first year (CHF 40.7 million in 2023), assuming that all eligible persons would be contacted in 2023. From 2024 onwards, risk assessment was limited to persons newly achieving the minimum age of eligibility (i.e., persons reaching 60 years of age). As a consequence, the costs for risk assessments decreased to CHF 2-3 million per year. Second, the large amount of people identified as eligible for screening due to the risk assessment in the first year automatically led to higher number of identified lung cancer cases and subsequent treatments. Therefore, initial care costs in 2023 and 2024 were much higher than in subsequent years.

After high initial costs due to risk assessment and initial care for newly identified lung cancer cases, the cost difference between screening and no screening decreases significantly (from CHF 141.0 million in 2023 to CHF 40.9 million in 2030 and CHF 21.6 million in 2037). This is also due to a strong decrease in the terminal care costs due to screening: While the costs for initial care and continuous care in the screening scenario are always higher than in the no screening situation (mainly due to the higher number of identified cases), the costs related to terminal care are only initially higher for screening. Thereafter, they are considerably higher for the no screening scenario (up to CHF 28 million higher than for screening). This is due to the stage shift of the identified lung cancers (i.e., due to screening, lung cancer is diagnosed at earlier stages at which a wider range of curative treatment options are available) leading to higher probability of remission (and lower probabilities to progress to advanced stages requiring terminal care).

Over a period of 15 years (2023-2037), lung cancer screening according to the above-mentioned criteria would costs up to CHF 810 million more than no screening. Assuming that all costs would be equally distributed over time, this would mean CHF 54 million per year.

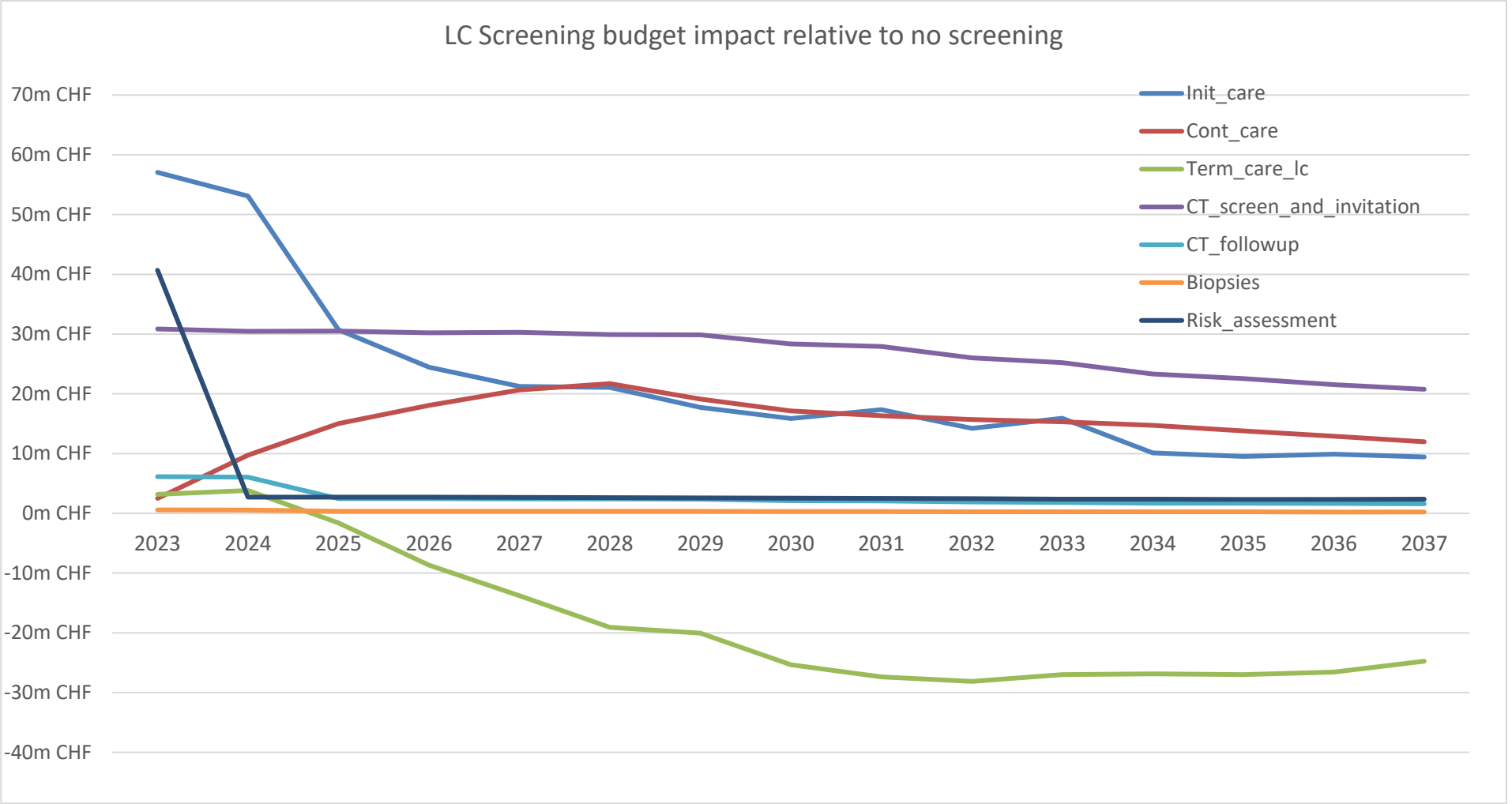


Figure 37 Budget impact relative to no screening for the first scenario on the efficiency frontier based on NLST inclusion criteria (start age 60 years, stop age 75 years, 40 pack-years, max 10 years since smoking cessation, biennial screening)

Year	Costs (in million CHF)							
	Initial care	Continuous care	Terminal care	CT screen and invitation	CT follow-up	Biopsies	Risk assessment	Total
2023	57.06	2.51	3.18	30.85	6.14	0.57	40.68	140.99
2024	53.08	9.73	3.82	30.47	6.06	0.57	2.72	106.46
2025	30.72	15.03	-1.59	30.50	2.45	0.34	2.69	80.14
2026	24.46	18.07	-8.66	30.20	2.44	0.33	2.69	69.53
2027	21.26	20.67	-13.80	30.31	2.42	0.33	2.66	63.84
2028	21.08	21.71	-19.10	29.93	2.41	0.33	2.63	58.99
2029	17.73	19.14	-20.04	29.89	2.37	0.33	2.59	52.02
2030	15.85	17.14	-25.34	28.33	2.12	0.30	2.53	40.95
2031	17.36	16.33	-27.38	27.91	2.06	0.30	2.51	39.08
2032	14.23	15.69	-28.13	26.01	1.91	0.28	2.44	32.43
2033	15.93	15.31	-27.01	25.22	1.82	0.26	2.38	33.92
2034	10.11	14.73	-26.87	23.33	1.71	0.25	2.37	25.63
2035	9.50	13.81	-27.01	22.53	1.68	0.24	2.32	23.09
2036	9.92	12.92	-26.55	21.52	1.63	0.23	2.34	22.01
2037	9.43	11.99	-24.77	20.76	1.60	0.22	2.36	21.60

Table 34 Budget impact relative to no screening for the first scenario on the efficiency frontier based on NLST inclusion criteria (start age 60 years, stop age 75 years, 40 pack-years, max 10 years since smoking cessation, biennial screening)

18.2.4. Scenario based on NELSON criteria: Screening age range 60-80 years, at least 15 cig/day for 30 years or 10 cig/day for 35 years, max 10 years since smoking cessation, annual screening

Compared to the first scenario on the efficiency frontier based on NLST selection criteria, in this scenario the percentage of the population that was considered ever eligible for screening during their lifetime was higher (11.2%, ca. 457,000 persons born between 1940 and 1979). Figure 38 and Table 35 show the budget impact of this scenario compared to no screening. The graphical representation is very similar to the previous budget impact example. However, it should be noted that the costs for this scenario were much higher. The cost difference between screening and no screening was estimated to range between CHF 378.8 million in 2023 and CHF 103.7 million in 2037. Over a period of 15 years (2023-2037), lung cancer screening according to the above-mentioned criteria would costs up to CHF 2.400 billion more than no screening. Assuming that all costs would be equally distributed over time, this would mean CHF 160 million per year.

It should be remembered that scenarios adopting broader eligibility criteria (i.e., larger age range, lower smoking intensity, and lower risk) are more expensive but also lead to higher benefits in terms of lung cancer death avoided and QALY gained (see cost-effectiveness section).

18.2.5. Scenario based on PLCO criteria: Screening range 55-80 years, PLCOm2012 risk threshold 1.7%, annual screening

Compared to the previous scenarios based on NLST and NELSON criteria, this scenario considered an even higher percentage of the population as ever eligible for screening during their lifetime (17.6%, ca. 720,000 persons born between 1940 and 1979). Again, the budget impact of this scenario looks similar to the previous ones (see Figure 39). As reported in Table 36, the cost difference between screening and no screening was estimated to range between CHF 485.4 million in 2023 and CHF 145.8 million in 2037. Over a period of 15 years (2023-2037), lung cancer screening according to the above-mentioned criteria would costs up to CHF 3.228 billion more than no screening. Assuming that all costs would be equally distributed over time, this would mean CHF 215 million per year.

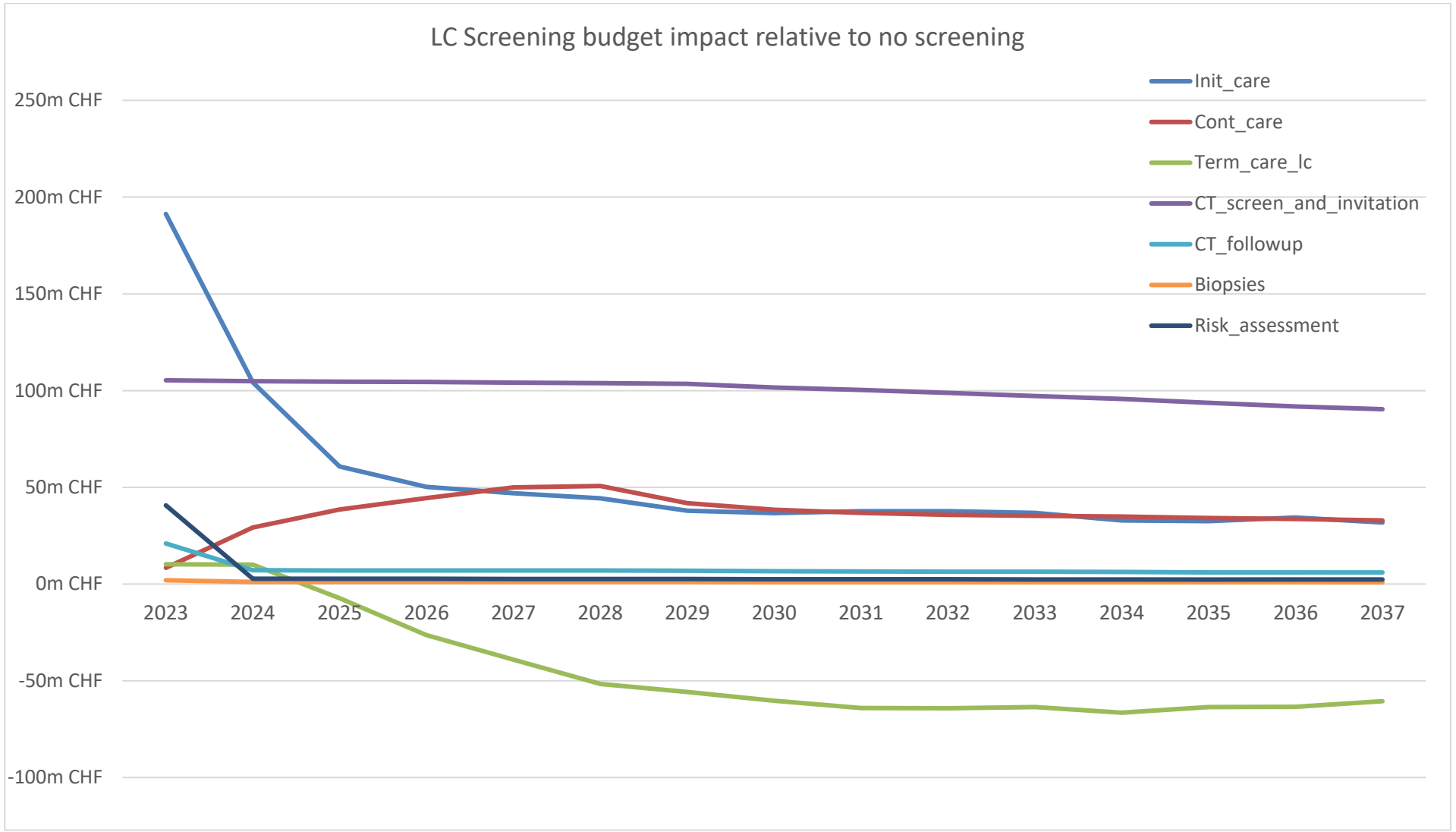


Figure 38 Budget impact relative to no screening for a scenario based on NELSON inclusion criteria (start age 60 years, stop age 80 years, at least 15 cig/day for 30 years or 10 cig/day for 35 years, max 10 years since smoking cessation, annual screening)

Year	Costs (in million CHF)							
	Initial care	Continuous care	Terminal care	CT screen and invitation	CT follow-up	Biopsies	Risk assessment	Total
2023	191.30	8.37	10.23	105.31	20.96	1.95	40.68	378.80
2024	104.17	29.28	10.00	104.85	7.16	1.08	2.72	259.26
2025	60.75	38.59	-7.30	104.61	7.04	1.07	2.69	207.45
2026	50.16	44.46	-26.42	104.47	7.03	1.07	2.69	183.45
2027	47.00	49.98	-39.06	104.14	6.98	1.07	2.66	172.77
2028	44.35	50.69	-51.63	103.88	6.97	1.06	2.63	157.96
2029	37.88	41.83	-55.81	103.41	6.93	1.06	2.59	137.90
2030	36.59	38.43	-60.32	101.62	6.62	1.03	2.53	126.50
2031	37.71	36.82	-64.02	100.27	6.54	1.01	2.51	120.84
2032	37.67	35.83	-64.19	98.79	6.43	1.00	2.44	117.98
2033	36.79	35.21	-63.60	97.22	6.32	0.98	2.38	115.31
2034	32.89	34.91	-66.48	95.65	6.24	0.97	2.37	106.54
2035	32.53	34.14	-63.61	93.62	6.05	0.94	2.32	106.01
2036	34.45	33.62	-63.40	91.76	5.98	0.93	2.34	105.68
2037	31.81	32.83	-60.60	90.38	5.98	0.92	2.36	103.67

Table 35 Budget impact relative to no screening for a scenario based on NELSON inclusion criteria (start age 60 years, stop age 80 years, at least 15 cig/day for 30 years or 10 cig/day for 35 years, max 10 years since smoking cessation, annual screening)

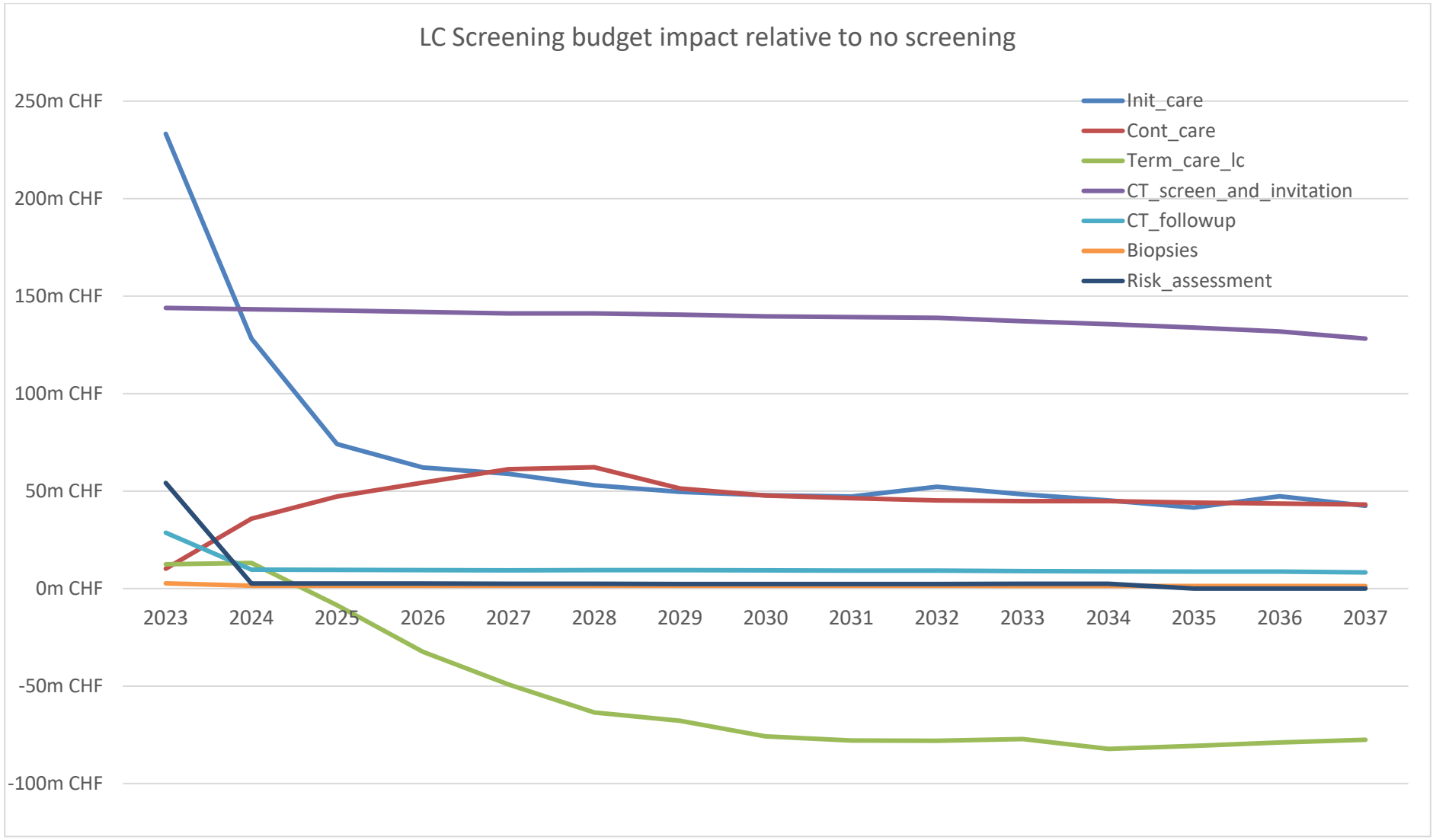


Figure 39 Budget impact relative to no screening for a scenario based on PLCOm2012 risk assessment (start age 55 years, stop age 80 years, risk threshold 1.7%, annual screening)

Year	Costs (in million CHF)							
	Initial care	Continuous care	Terminal care	CT screen and invitation	CT follow-up	Biopsies	Risk assessment	Total
2023	233.24	10.18	12.49	143.94	28.65	2.68	54.18	485.36
2024	128.29	35.81	13.14	143.18	9.68	1.47	2.62	334.21
2025	74.16	47.19	-8.53	142.57	9.53	1.46	2.56	268.94
2026	62.12	54.36	-32.40	141.92	9.43	1.45	2.54	239.42
2027	58.84	61.27	-49.13	141.14	9.38	1.44	2.47	225.41
2028	53.00	62.22	-63.61	141.07	9.47	1.45	2.40	206.00
2029	49.59	51.39	-67.75	140.50	9.40	1.44	2.39	186.94
2030	47.84	47.76	-75.82	139.59	9.27	1.42	2.35	172.42
2031	47.26	46.28	-77.93	139.20	9.22	1.42	2.36	167.81
2032	52.26	45.16	-78.11	138.89	9.19	1.41	2.38	171.19
2033	48.32	44.87	-77.23	137.18	8.93	1.39	2.40	165.85
2034	45.16	44.89	-82.22	135.56	8.85	1.37	2.43	156.03
2035	41.53	44.10	-80.64	133.90	8.77	1.36	- *	149.01
2036	47.35	43.62	-78.98	131.85	8.67	1.34	- *	153.84
2037	42.41	43.10	-77.53	128.21	8.28	1.29	- *	145.76

Table 36 Budget impact relative to no screening for a scenario based on PLCOm2012 risk assessment (start age 55 years, stop age 80 years, risk threshold 1.7%, annual screening)

*The risk assessment costs are applied to 25% of the population reaching the minimum age for eligibility. For screening start age at 55 years, it is expected that all persons will reach the minimum eligible age range as early as 2034, at which point we assume the requisite individuals have been risk-assessed

18.3. Discussion

In the absence of screening, the total costs related to lung cancer treatment in Switzerland for the 1940-1979 cohort was estimated to increase from CHF 474 in 2023 millions to CHF 724 millions in 2037.

Compared to no screening, the overall costs of all screening scenarios were generally higher. This was especially true in the first years in which we assumed that most of the risk assessments would happen (leading to costs for risk assessment and increased costs for the initial care of newly identified lung cancer cases). Over a period of 15 years, the total costs of lung cancer in the no screening scenario were estimated to reach CHF 9.4 billions, while the costs for the three highlighted scenarios on the efficiency frontier would range between CHF 10.2 billions and CHF 12.6 billions (i.e., +9% and +34% compared to no screening, respectively).

It should be emphasized that the screening scenario modelled in the cost-effectiveness analysis, including those on the efficiency frontier, varied considerably in terms of selection criteria (i.e., screening age range, smoking intensity, risk threshold). This resulted in a large variation of the number of persons considered ever eligible for screening during their lifetime, which ranged between 7% and more than 20% of the population born between 1940 and 1979. The number of persons screened as well as the number of lung cancers identified (and treated) have a big influence on the budget impact of screening. Therefore, screening eligibility criteria play a major role on the estimated budget impact.

Another important remark concerns the participation rate of a lung cancer screening program. The current analyses assumed that all persons considered eligible for screening would participate. However, the actual participation rate may be much lower, and the real budget impact of the above-mentioned screening scenarios may be considerably different. A recent article published by Baldwin et al. nicely summarized the current evidence on participation to lung cancer screening: "*The evidence from clinical trials in lung cancer screening is limited because only one has used a true population approach in recruitment, the United Kingdom Lung Screen pilot trial (UKLS), and offering screening as part of a clinical trial may underestimate the participation rates of a real-world service. In the UKLS trial, 31% of eligible people responded to an initial questionnaire but only 11.5% of participants were at high enough risk for trial entry and 47% of these gave their consent. In the NELSON trial a population approach was initially used for adult males only. Thirty-two percent of those eligible responded to a questionnaire on general health, lifestyle, and smoking history. Nineteen percent of the respondents met the eligibility criteria for the trial and received an invitation for participation in the trial, an information leaflet, and an informed consent form combined with a short questionnaire. Of these individuals, 51% gave informed consent and were recruited… In the US, where LDCT screening has been funded since 2015, participation rates were 3.3% of the eligible population in 2015 and more recently*

estimated to be 14% in 2018 although only 4% in the uninsured."[183]. It should be emphasized that these rates are based on clinical trials, pilot studies, or very early participation rates. Therefore, it may be more appropriate to look at participation rates to screening for other cancers in Switzerland. For example, a recently published report on breast cancer screening in Switzerland reported that of the 910,485 (half of 1,820,969) eligible women in 2016-2018, 774,348 were invited to breast cancer screening (coverage by invitation rate of 85.0%). The participation rate was estimated to be 45.8% [184]. Another study investigating the attendance to cervical cancer screening in Switzerland using data from the Swiss Health Survey 2012 reported a participation rate of 72.9% for the 20-69 year-old women living in Switzerland within the last three years.[185] Whether a lung cancer screening program may reach similar participation rates is unclear and may depends on many factors (e.g., screening organisation, adopted eligibility criteria, available infrastructure and medical personnel).

19.Ethical issues in lung cancer screening

19.1. Introduction and methodology

Screening for lung cancer raises a variety of complex and intertwined ethical issues. This chapter describes and analyses these issues in order to inform the conclusions of the wider health technology assessment. The normative analysis provided here draws on the empirical evidence regarding clinical effectiveness and cost-effectiveness provided in the literature and in the preceding parts of this report, as well as the available evidence on ethical issues. In this part of the report, the relevant evidence is referred to in relation to each ethical issue, rather than a in separate literature section, as this enables more nuanced analysis and avoids repetition.

Literature was identified using purposive sampling on PubMed and Google Scholar, followed by further identification of relevant references and screening of abstracts. Papers were selected if they referred to ethical issues relating to lung cancer screening or patient attitudes to screening, but only those that dealt with such issues in depth were ultimately included in the review. The only new ethical (sub-)issue identified in the literature review but not anticipated in the ethics section of the scoping report was the potential for bias in committees involved in screening decisions (see section 19.5.1.). One other ethical issue regarding a possible increase in stigmatisation in relation to implementing screening was identified during the ethical analysis (see section 19.5.2.).

Though there are many thematic connections between them, the ethical issues concerning screening fall into two broad categories. Several ethical issues are raised by the use of screening in the clinical context. First, those eligible for screening may be reluctant to engage in it, and there is an associated possibility of stigmatisation; attitudes also differ among current smokers, ex-smokers, and different demographic groups. Second, perhaps the most important issue concerns the shared decision-making process: Potential participants in screening must be fully informed and engage in shared decision making with physicians and other healthcare professionals before reaching a decision about whether to engage in screening. This is particularly important, as generally speaking, both members of the public and some healthcare professionals can overestimate the benefits of various types of screening, with bias towards perceiving benefit rather than potential risks and costs. [186] Third, there are various ethical issues relating to screening modalities, such as the costs to participants in terms of false positive results, travel, number of screenings attended and radiation exposure, and incidental findings. Finally, there is also the issue of whether screening should be offered in conjunction with smoking cessation.

The second category of ethical issues is wider concerns of justice and fairness. Here, the first and most fundamental issue concerns the ethical use of resources: Do the benefits of lung cancer screening justify the substantial cost of implementing such a programme? Judgements about cost-effectiveness can be subjective, and may ultimately involve values as well as evidence. Second, even if screening is deemed cost-effective in a given health care setting, there may be concerns about perceptions of moral responsibility for lung cancer, or for increasing one's risk of lung cancer. Specifically, some members of the public may view it as unfair to provide screening for what is perceived (by some) as a "self-inflicted" disease; in particular, it might seem unethical to offer screening to a smoker who is making no effort to quit (even if this perception is inaccurate). Third, in addition to concerns about distributive justice (fair use of resources), implementing a screening programme may pose challenges for societal justice by exacerbating health inequalities through differential effects on certain groups in society; for example, people in lower socioeconomic groups may benefit more from screening but are also less likely to engage in it, meaning that disparities could increase if those in higher socioeconomic groups participate more in screening. Finally, issues of justice and fairness are also raised by the eligibility criteria for screening; for example, if screening is restricted to smokers and ex-smokers, this could exclude other groups who are at higher risk than the general population, including those who are at increased risk of lung cancer through occupational exposure to dangerous particles such as asbestos, those who have never smoked but have had heavy exposure to passive smoke, or those who have a family history of lung cancer.

As already noted, many of these issues are inter-related. The following sections proceed through them in the order set out above, while noting any important connections between ethical issues.

19.2. Clinical ethical issues

19.2.1. Public attitudes to screening

Before proceeding to the specific factors that can act as barriers to participation in screening, an overview of public attitudes to screening will provide useful context. The literature shows that there is broad support for screening among smokers, ex-smokers and never smokers. A study in England found that 97% of the public thought that screening was a good idea.[187] Another study in Wales found that over 90% of respondents believed screening conferred benefit. [188] In a Belgian study the figure was lower, at 83.6%; the same study found that 84.3% of current and ex-smokers would participate in a screening programme. [189]

Interestingly, one study showed greater support for screening among never smokers than among smokers, which may appear counterintuitive given the assumption that non-smokers would 'blame' smokers and not support screening (see section 19.5.2. for more on this issue[187] [190]). Similarly, a US

study found that only 71.2% of smokers vs 87.6% of never smokers would consider lung cancer screening. [191] This is also somewhat ironic, given that never smokers are highly unlikely to benefit from screening and smokers stand to benefit most. However, these findings may be explained by some smokers' negative attitudes towards screening. While smokers do tend to support screening, this group is also more likely to be reluctant to engage in screening or see it as not worthwhile. A qualitative study found that many smokers were fatalistic about lung cancer and were often unaware of the potentially curative treatment options; they were also more likely to be worried about being called back repeatedly for further scans. [191] An association has been identified between not wanting to engage in screening and three specific factors: "Fatalism, low perceived value of symptom presentation, and having negative views about treatment."[188]

The relative lack of enthusiasm for screening can also extend to interventions following a positive screening result: The same US study found that "Only half of the current smokers would opt for surgery for a screen-diagnosed cancer."[191] Again, negative views about treatment for detected cancer could explain this result; if half of current smokers would refuse treatment for lung cancer, there might seem little point in engaging in screening in the first place. Interestingly, and in contrast to pessimism about treatment, almost 30% of current smokers in one study agreed with the statement that "If the CT scan is negative, you can continue to smoke without worrying about lung cancer."[190] If close to a third of smokers believe this, then screening could undermine smoking cessation interventions; furthermore, it suggests that smokers are misinformed to a large extent.

Stigma is also an important aspect of smokers' attitudes to lung cancer screening. A systematic review of 7 qualitative and 8 quantitative studies found that the general stigma attached to lung cancer also affects screening for the condition: "Patients reported feeling stigmatised by the prevailing view that if someone had lung cancer they would necessarily be a smoker and have inflicted this disease on themselves; and this view was seen by patients as unfair. Patients feared that they would be denied treatment and thought that lung cancer was neglected in research and screening because of the link between smoking and lung cancer."[192] Physicians and other healthcare professionals must be alert to the risk of stigmatisation and take care to avoid giving any inadvertent impression that smokers should be blamed for their lung cancer, or for increasing the risk of developing the condition.

Indeed, all of these factors, including reluctance, stigmatisation, misunderstanding and fatalism, must be borne in mind when inviting, discussing and making decisions about lung cancer screening with patients (see section 19.2.2). It is also important to bear in mind that attitudes vary among different demographic groups. For example, people from lower socioeconomic groups and minority ethnic groups are more likely to develop lung cancer and more likely to have difficulty in accessing and

adhering to screening, [193] and demographics also affect attitudes to screening. Section 19.5.3. explores these issues in more detail.

19.2.2. Shared decision making and risk communication

Potential participants in lung cancer screening must decide whether to take part based on discussion with their physician and/or healthcare team. In contrast with the old paternalistic model, where physicians tended to make decisions on patients' behalf without much regard for their views, the increased emphasis on respect for patient autonomy has led to great importance being accorded to the process of obtaining informed consent from the patient.[194] Older models of informed consent relied simply on physicians providing information and the patient making a choice based on that information, but the more modern paradigm of shared decision making uses a more cooperative model, where the decision is the ultimate product of discourse and discussion between physician and patient. The United Kingdom's National Institute for Health and Care Excellence (NICE) defines shared decision making as "A collaborative process that involves a person and their healthcare professional working together …. choosing tests and treatments based both on evidence and on the person's individual preferences, beliefs and values." [195].

NICE recently published guidelines on shared decision making that are very helpful for consideration of ethical issues in lung cancer screening. The guidelines state that patients should be helped to "Actively engage in the discussion; explain what matters to them; make decisions about their care; [and] remember information they have been given during the discussion" and that including family members in the discussion can help achieve these objectives.[195] More specifically, the healthcare professional should:

- explain the healthcare aims of each option and discuss how they align with the person's aims, priorities and wider goals
- openly discuss the risks, benefits and consequences of each option, making sure the person knows this includes choosing no treatment, or no change to what they are currently doing
- clarify what the person hopes to gain from a treatment or intervention and discuss any misconceptions
- set aside enough time to answer questions, and ask the person if they would like a further opportunity to discuss options.[195]

As the second point here suggests, the issue of risks versus benefits lies at the core of the discussion about lung cancer screening. Many patients and some clinicians overestimate the potential benefits of screening while underestimating potential harms, and it is essential that information provided to patients is objective and understandable to avoid addition of any biasing factors. Furthermore, if

patients bring their own biases or misunderstandings to the discussion about screening, these should also be addressed as part of that discussion; if the physician has any particular values or biases that he or she believes might influence the decision, these too should also be shared with the patient to enable transparent dialogue. [196]

The prospective benefit of lung cancer screening is, of course, that it can increase the chance of a tumour being detected early, and thus also increase the chance of earlier treatment and avoidance of or delay in death due to lung cancer. A second benefit is potential peace of mind for those screened. The potential harms and burdens are less obvious to the public. First, there is a risk of false positive results, where cancer is detected but is not actually present; such false positives can result in repeat scans and unnecessary surgery.[197] A related concern is overdiagnosis, where cancer is correctly detected but might never have caused issues for the patient; here, too, unnecessary scans, surgery and treatment can result.[197] Second, although low-dose CT screening involves less radiation, patients are still exposed to ionising radiation as part of screening. [197] Third, taking part in screening can increase anxiety among patients, though it can also lessen anxiety in some. [197] Fourth, there is a possibility that incidental findings (unanticipated scan results revealing an additional medical issue) might be reported to the patient with the consequences of additional diagnostic procedures. [197] Finally, attending screening involves a time and travel burden for participants.[193]

All of these prospective benefits and burdens must be discussed with patients in the clinical encounter concerning lung cancer screening. More specifically, the risks and benefits of screening must be communicated in a comprehensible way that patients can understand. The need to explain all these aspects to patients and answer questions means that the discussion may require a considerable amount of time – time that is not reimbursed by the health care system in the context of screening. The patients most likely to benefit from lung cancer screening belong to lower socioeconomic groups, and are thus also more likely to have lower levels of health literacy, and lower levels of education generally, so communicating with them about the risks and benefits of can be particularly challenging.[190]

Generally, it is best practice to personalise risk evaluations as much as possible for patients, but (in the absence of a precision medicine approach) this may not be possible in the case of lung cancer screening beyond providing statistics that relate to the person's smoking status (current/ex/never smoker). Vague terminology such as 'rare', 'unusual' and 'common' should be avoided, and concrete numerical descriptions used instead.[195] NICE provides useful specific guidance in this regard:

- Use absolute risk rather than relative risk. For example, the risk of an event increases from 1 in 1,000 to 2 in 1,000, rather than the risk of the event doubles.

- Use natural frequencies (for example, 10 in 100) rather than percentages (10%).
- Be consistent when using data. For example, use the same denominator when comparing risk: 7 in 100 for one risk and 20 in 100 for another, rather than 1 in 14 and 1 in 5.
- Present a risk over a defined period of time (months or years) if relevant. For example, if 100 people have treatment for 1 year, 10 will experience a given side effect.
- Use both positive and negative framing. For example, treatment will be successful for 97 out of 100 people and it will be unsuccessful for 3 out of 100 people. [195]

This advice concerns diagnosis, treatment and testing, but it will be helpful to apply it to discussions about lung cancer screening. How should the main risks and benefits of lung cancer screening be explained to patients? Using the conclusions of the preceding parts of this report, we can construct some example information for patients. Note, however, that the numbers used in discussion of harms and benefits here is primarily illustrative, in order to indicate how deliberations regarding screening should proceed. Given any substantial changes in the statistics, the ethical conclusions might also be altered.

The key finding from the clinical effectiveness part of this report was that "The risk ratio (RR) of death from lung cancer of LDCT compared with no screening or CXR in trials with ≥ 5 years of follow up was 0.80 (95%CI 0.72-0.88)." However, it is not advisable to communicate risk to patients using risk ratios like this; more helpful is the number of deaths with and without screening in a group of patients. This report puts that figure at 207 per 10,000 patients without screening, and 164 deaths with screening (Table 12). Most patients could understand this explanation, but it is not clear how informative it is for decision making, even if translated into percentages: after at least five years of follow up, 2.07% of patients die without screening, and 1.64% of patients die with screening. Translating the risk into natural frequencies may be more helpful; there is a 1 in 50 risk of dying of lung cancer without screening, and a 1 in 61 risk with screening. This is a substantial difference that many patients will find persuasive. Note that if communicated to the patient in terms of relative risk, contrary to the NICE guidance, this relative reduction would be expressed as 22% lower risk of death (61/50); in absolute terms, however, it remains a reduction of only 0.43%. This point emphasises the importance of using absolute rather than relative risk when communicating with patients.

However, as per the NICE guidance on risk communication, the time factor must also be considered. As stated in the clinical effectiveness part of this report, "If 1,000 individuals at high risk of lung cancer were to be screened at intervals of 1, 2, and 2.5 years and then followed for 10 years 139 individuals would have died from any cause in the screening and 138 would have died without screening but 7 additional individuals (in the best case 12 and in the worst case 2 individuals) in the LDCT group

compared to individuals with no screening would have been saved from death from lung cancer." This formulation makes it clear to potential participants in screening that screening that, while lung cancer mortality was reduced by participation in screening, overall mortality was not substantially affected. While the risk ratio was 0.96 with screening, this translates to only 36 fewer deaths per 10,000 (71 fewer to 0 more per 10,000; 95% CI, as stated in Table 12). In other words, screening might reduce a patient's chances of dying of lung cancer, but it is unlikely to prolong his or her life. This might be difficult to explain to patients, but should nonetheless be disclosed as part of the shared decision making process, given that avoiding or delaying death is likely to be a key motivation among patients.

Now we can turn to the literature regarding the main negative effect of screening, getting a false positive result, which can lead to negative psychological effects.[198] Here the evidence evaluated in this report is weaker, with false positives defined differently between trials and a range between 0.6% and 46%. More important from the patient perspective is the number of false positive scans with invasive procedures, which varied between 2.6% and 9.6% among those invited for follow-up. However, these numbers relate to a wide variety of trials under different conditions, and given that a pooled estimate for invasive procedures following false positives was not possible, we will adopt the estimate from the NELSON trial, as recommended in the clinical effectiveness part of this report, which states that NELSON found that "In 293 of 493 participants undergoing a work-up lung cancer was not confirmed (false positive rate of 1.2% (defined as [recall scans or work-ups – screening detected cancers] / total number of individuals)." Taking this figure (which also falls in the mid-range of comparable results from different studies), it is reasonable to conclude that patients could be informed that the risk of a false positive scan that results in unnecessary surgery is around 1.2%; in other words, around 1 in 80 patients will experience this harm over 10 years.

False negatives can also occur; these are defined as negative scans after which cancer is detected within a year. However, the clinical effectiveness part of this report found these rates to be very low, at less than 1% in most trials. The risk of overdiagnosis, where cancer is identified and treated but would not have posed any issues for the patient in his or her lifetime, meaning that treatment was unnecessary, must also be considered. In the NELSON trial, this was calculated to be 8.9% [61] over a decade of follow-up, suggesting that almost 1 in 10 identified cases of cancer were overdiagnosed.

How patients weigh up the potential benefit of reducing the risk of dying from lung cancer and the potential harms of a false positive or overdiagnosis is a matter of personal values, but the physician should always be prepared to help inform the shared decision making process. A 1 in 80 risk of surgery resulting from a false positive may seem a small price to pay for reducing the risk of dying of lung cancer by even a small amount; however, the negligible reduction in overall mortality should also be

shared with patients and may have a dissuasive effect. It seems likely that screening might have less of a beneficial effect on peace of mind if potential participants in screening were informed that, while it may reduce the risk of death from lung cancer, this may not actually lead to prolongation of life, as many would assume.

19.3. Screening modalities

Patients' decisions about whether to engage in lung cancer screening will depend not only on their assessment of the risks and benefits of screening itself; how the screening is implemented is also important. Factors such as the cost of screening, the location of screening, the frequency of screening, the handling of incidental findings and the potential linkage of screening with participation in smoking cessation programmes are not ethical issues as such, but each of them can raise ethical issues and this can complicate engagement with potential screenig participants.

In Switzerland, screening costs could be reimbursed by insurance providers in the event that screening is deemed to be effective and cost-effective, but the cost of screening tests is known to influence uptake of screening. For example, in a study in the United States, the percentage of smokers willing to pay $200 for a test was only 27%, half of the proportion of never smokers; when the cost was increased to $300 only 11% of smokers and 27% of never smokers would pay.[191] Regardless of how a screening programme is funded, it is important to discuss potential costs or lack of them when discussing participation with patients in order to avoid deterring them.

The location where screening is provided also plays a role in decision-making, with one study finding that it is an important factor for three quarters of potential participants; hospital-based programmes are less popular than community-based programmes, with almost 1 in 4 saying they would not attend a hospital programme. [199] This trend was more pronounced in smokers and those from deprived communities. Among those who were unlikely to attend a hospital programme, over 80% cited travel being a potential obstacle, indicating the perceived advantages of community based screening programmes. Notably, over twice as many African-American as white participants cited convenience and cost as major concerns, indicating the substantial association between demographic group and attitudes to screening. [200]

The issue of frequency of screening relates closely to the previous two issues; the more frequent the screening, the higher the cost and the potential inconvenience of travel. Furthermore, more frequent screening could increase the psychological burden on patients, as well as increasing their exposure to radiation. Recent studies suggest that annual screening is not necessary for all patients, and that a gap of 1-2 years may be optimal. [201] Given that the optimal interval may vary for different patients, this issue should also be discussed with them.

The possibility of incidental findings being detected on an LDCT screen must be made clear to patients as part of the shared decision making process. A study in England found that 10% of participants in screening were referred to primary care for suspected undiagnosed chronic obstructive pulmonary disease or cardiovascular disease risk assessment. [202] Ultimately, patient management changed only for a fifth of those referred, but the potential added inconvenience of incidental findings could be important for some patients, even if they represent a potential added benefit in terms of prevention of future disease. The way in which incidental findings are handled may vary between screening programmes (particularly between hospital and community programmes), and the process for handling them if they arise should also be explained to patients.

19.4. Smoking cessation

Smoking cessation is known to improve lung cancer outcomes, and it might be thought that participation in cessation programmes should be a condition of lung cancer screening (see section 19.5.2 for more on this and related issues relating to personal responsibility). However, the evidence suggests that making smoking cessation mandatory would deter smokers from participation in screening programmes. One study found that over 70% of smokers were willing to receive cessation advice at the same time as screening, but that it would deter almost 10% from participation in screening.[189] Smokers should of course be advised in general terms that cessation is likely to improve health outcomes, but the potential advantages of providing cessation advice concurrently for the majority of smokers must be weighed against the risk of deterring from screening those least likely to engage in cessation programmes. In any case, it is clear that participation in cessation counselling should not be a mandatory criterion for participation in screening.

19.5. Justice, fairness and cost-effectiveness

19.5.1. Cost-effectiveness and justice

Having discussed the many ethical issues regarding provision of lung cancer screening to patients, we now adopt a wider perspective and consider whether screening represents just and fair distribution of healthcare resources and funding. Even if screening were found to benefit many patients, it remains possible that the cost of providing screening is so great that it is not cost-effective – or indeed just - to provide it.

Importantly, these cost concerns are not only financial but also medical. In order for one patient to benefit from screening, many patients must be screened; they bear the potential burdens of screening in the hope that they benefit, even though they may not. In many ways, deliberation regarding whether lung cancer screening represents fair distribution of resources parallels the individual patient's consideration of whether the benefits of screening outweigh the potential harms and

burdens. As such, just as the judgement regarding whether screening is cost-effective depends on which threshold is set for cost-effectiveness, the question of whether screening represents a just use of resources is rather subjective and depends on what the decision-maker regards as an appropriate distribution of resources within the healthcare system. Notably, the literature suggests that this subjectivity can affect not only patients and physicians, but also committees and other organisations making decisions about screening programmes; this means that "Committees should explain their policy decisions with reference to values as well as evidence, so that values considered in decision-making can be interrogated and challenged if necessary." [203]

The following paragraphs explore cost-effectiveness and justice of screening provision using the data provided in the previous chapters of this report. It should be noted from the outset that cost effectiveness and justice are clearly not synonymous phrases: Even if a particular screening programme is cost-effective according to health economics, it may not be just to implement it, depending on which resources are available and what other programmes are implemented. In addition, what exactly is meant by "cost-effective" is of prime importance.

The conclusion of the health economics part of this report was that "Several lung cancer screening strategies may be cost-effective in Switzerland." More specifically, it found that the potential cost per QALY gained compared to no screening ranges between CHF 14,452 and CHF 37,959. These are wide ranges, and the final assessment regarding cost-effectiveness and justice in terms of the distribution of resources may change according to where the most accurate figures lie within these ranges.

However, the conclusion regarding cost-effectiveness may also vary substantially even given relatively similar models, as the literature on cost-effectiveness reflects. The recent UKLS trial found that, with 1,987 participants receiving screening and 1,981 in the control arm, 30 lung cancer deaths were reported in the screening arm, and 46 in the control arm, with meta-analysis of this and other trials revealing an overall relative risk of mortality from lung cancer of 0.84; however, the reduction in all-cause mortality was less, at only 0.97. Nonetheless, the authors concluded that "even a small reduction in all-cause mortality as shown here, does represent a large number of lives should countries around the world adopt lung cancer screening programmes."[70] This conclusion rests on the assumption that results from screening trials are externally valid and thus transferable to the real world. In any case, it is true that this would represent a large number of lives saved; however, it would also represent a huge number of screenings that yield no clinical benefits for most patients, as well as vast expenditure of resources. The optimistic conclusion of the UKLS authors provides an interesting contrast with the more pessimistic judgement of an earlier modelling study of CT:[204]

"Under these ideal circumstances, the absolute reduction in lung cancer mortality was 900 people per 100 000, a 16 percent relative difference. The number harmed by unnecessary tests increased to 1520 per 100 000, and the cost per QALY gained was $42 500. Quitting smokers and former smokers had adjusted costs of $75 300 and $94 400 per QALY gained, respectively. Even under the most favorable of circumstances, CT scanning at current cost per scan seems unlikely to be highly cost-effective as a screening test for lung cancer." [205]

Here, even though the reduction in lung cancer mortality given a best-case scenario model was exactly the same as in UKLS, the authors concluded the opposite, stating that screening is unlikely to be cost-effective (and this despite lower rates of unnecessary tests under this scenario than under the baseline scenario). This study was conducted in 2003, and screening has clearly made great advances since then, but nonetheless the subjectivity involved in such decisions is clear given the vastly differing interpretation of quite similar statistics. Whether a programme is deemed cost-effective will also depend on which other similar (and dissimilar) screening programmes are implemented in a given country or healthcare system (see below).

As stated in section 7.6, the conclusion of the clinical effectiveness part of this report was that there are 207 lung cancer deaths per 10,000 patients without screening, and 164 deaths with screening. Weighing those risks is a subjective decision that must ultimately be dealt with at the individual patient level, but we can also use these data to inform our discussion of justice and fairness in relation to screening.

It may be just to try to prevent 43 lung cancer deaths by screening 10,000 people, with 120 of them having unnecessary surgery as a result. However, from the perspective of justice, we can combine these numbers in a more informative way that enables comparison with other forms of screening. If screening 10,000 people is necessary to prevent 43 deaths from lung cancer, with 120 having unnecessary surgery, that means that approximately 3 people will undergo the harm of unneeded surgery to prevent one death from lung cancer. (This comparison combines the conclusions regarding effect on mortality of LDCT from this report with the false positive/invasive procedure rate from NELSON; using NELSON's own effectiveness figures, this figure was estimated to be 130 people screened and 33 unnecessary biopsies for one death prevented over 10 years of follow-up.[206]) Furthermore, of those who have cancer correctly detected by screening, perhaps 10% will be overdiagnosed, meaning that they undergo further interventions unnecessarily.

For the sake of illustration, comparison with a different form of screening will be helpful. First, a comparison of effectiveness. When the Swiss Medical Board conducted an evidence review of breast cancer screening, they concluded that screening 1,000 women would prevent 1 death from breast

cancer, though there was no apparent impact on overall mortality. In addition, “for every breast-cancer death prevented in U.S. women over a 10-year course of annual screening beginning at 50 years of age, 490 to 670 women are likely to have a false positive mammogram with repeat examination; 70 to 100, an unnecessary biopsy; and 3 to 14, an overdiagnosed breast cancer that would never have become clinically apparent.” [207] The Board also pointed out that women’s peceptions of the benefits of screening are overoptimistic, with the mistaken impression that screening reduces breast cancer mortality by 50%. These conclusions were controversial and have since been contested, but they provide a useful comparator of a case where it was deemed unethical to continue with screening on the grounds of the distribution of harms and benefits. In comparison, for LDCT lung cancer screening, one study [208] in the United States found that “for every 1000 people screened, 10 will be diagnosed with early-stage lung cancer (potentially curable), and 5 with advanced-stage lung cancer (incurable); 20 will undergo unnecessary invasive procedures (bronchoscopy and thoracotomy) directly related to the screening; and 550 will experience unnecessary alarm and repeated CT scanning (with its associated irradiation)." [209] Thus, in this comparison, lung cancer screening compares favourably with breast cancer screening, with potentially, 10 times as many deaths from cancer prevented, and 3.5-5 times fewer invasive procedures.

Next, a brief comparison of lung and breast cancer screening in terms of cost-effectiveness. A 2007 study found that breast cancer screening programmes in Switzerland would be cost-effective, with a range compared with opportunistic screening of $73,018 to $118,193 – considerably more expensive than lung cancer screening. [210] This suggests that LDCT may be both more effective and more cost-effective than breast cancer screening. However, while LDCT may prevent more deaths and have fewer invasive false positives per person screened than breast cancer screening and thus enjoy a more favourable verdict in terms of cost-effectiveness, for both types of screening the impact on overall mortality appears to be negligible, and the Swiss Medical Board recommended discontinuing breast cancer screening on the basis of its evidence assessment.

These examples show that whether implementing lung cancer screening represents just distribution of resources depends not only on the costs and cost-effectiveness of that screening programme; it also depends on which other cancer screening programmes are provided. If, for example, lung cancer screening cost substantially less per LYG than breast cancer screening, yet was not implemented, this would correctly be perceived as unjust and potentially discriminatory against smokers and former smokers. If, in contrast, lung cancer screening costs substantially more per LYG than breast cancer screening, it might be unfair to implement such a programme given that this might mean other screening programmes or health interventions would thus have to go unfunded.

For the purposes of deliberations concerning justice, it is also important to consider precisely what is meant by the "effectiveness" part of "cost-effectiveness". LDCT screening is clearly effective at preventing deaths from lung cancer and cost-effectiveness calculations are focused on this key metric and the number of LYG: As stated in section 11.2.2 of this report, "the number of lung cancer deaths prevented with screening would range between 203 per 100,000 persons for the first scenario on the efficiency frontier to 1,071 per 100,000 persons for the last scenario on the efficiency frontier." In terms of cost-effectiveness, "The number of LDCT screen per lung cancer death avoided would range between 155 and 434 LDCT screens per LYG." Effectiveness here is thus measured, as might be expected, in terms of avoiding death from lung cancer. In these terms, lung cancer screening can clearly be cost-effective (depending on the threshold).

However, as stated in the clinical effectiveness section of this report, screening has little impact on overall mortality. It is a truism that an intervention can only be cost-effective if it is effective (Raftery et al [211]); if screening does not improve overall mortality, then it cannot be (or be expected to be) cost-effective in that particular regard. In terms of justice, the question is thus whether it is fair distribution of resources to spend potentially billions of Swiss francs over several decades given the negligible impact of LDCT on overall mortality. Preventing deaths from lung cancer is a laudable aim, but if participants in screening do not tend to live longer as a result, this may not represent the fairest use of resources even if screening is deemed cost-effective in terms of reducing deaths from lung cancer. (It should be noted that it would be inaccurate to depict this as a clash between opposing principles of justice and utility. [212] Rather, it is a question of being clear about exactly what the derived utility is before determining whether the cost paid to gain that utility represents just use of resources.)

In effect, there are two levels of fairness involved in discussing justice here. As stated above, it would be perceived as unfair not to implement lung cancer screening if it is more cost-effective than breast cancer screening, according to the paradigm of health economics (fairness in terms of screening programmes). However, from the wider perspective of justice it remains unclear whether the cost per LYG in terms of avoiding death from cancer is fair given the negligible effect on overall mortality (fairness across the healthcare system). Of course, this criticism may also apply to other cancer screening programmes, and it may be that a focus on preventing cancer deaths is more important for both healthcare system and society than a focus on reducing overall mortality.

Even if it is cost-effective in terms of preventing lung cancer deaths, it is not clear whether screening represents the fairest distribution of healthcare resources, particularly given the additional option of smoking cessation for current smokers, who constitute the most at-risk group. While screening can be effective in terms of the potential benefit of a reduction in lung cancer mortality, this benefit will not

accrue to most of those screened; more importantly, however, the overall reduction in mortality is marginal at best, and indeed the clinical effectiveness part of this report concluded that "overall mortality appears not to be affected by LDCT screening." If that is the case, it means that the expenditure of vast resources and the imposition of substantial inconvenience and harms for the screened population becomes difficult to justify. However, if other screening programmes with similar or lower cost-effectiveness and reductions in mortality are funded, then lung cancer screening could also be funded, assuming sufficient resources are available [213] and that it is agreed that reducing lung cancer deaths rather than reducing overall mortality is the appropriate aim.

19.5.2. Stigmatisation and perceived responsibility for lung cancer risk

As stated above in section 19.5.1., some smokers fear stigmatisation to the extent that it can deter them from screening. A related concern is that some members of the public might object to publicly or insurance-funded lung cancer screening on the grounds that lung cancer is perceived to be a "self-inflicted disease" which was easily avoidable and should not have (more) resources wasted on it. Even if this perception of lung cancer were accurate, of course, the conclusion that screening would be a waste of resources does not necessarily follow.

In any case, the available evidence suggests that the view that smokers should be "blamed" is not widely shared by the public, with many studies indicating similar levels of support among never smokers to those who smoked or had quit. A quantitative representative survey in England found that over 90% of respondents believed that screening improves survival, with no association between smoking status and support for lung cancer screening, meaning that non-smokers supported screening just as much as smokers and ex-smokers (this finding also applied to specific negative and positive attitudes). A Belgian study found that 84% of respondents supported screening, with no significant difference between current/ex-smokers and never smokers.[189]

Another English study found that 97% of respondents thought that screening was a good idea, but also that over a fifth thought that screening was a waste of National Health Service (NHS) funds. Specifically, "using NHS money to screen smokers was perceived as a waste of NHS money by 21%, but most commonly by former (24%) and never smokers (22%) compared with current smokers (14%)."[187] The reasons or justifications underlying these negative attitudes were not sought in the survey, but the authors conclude that "the stigma attached to smoking may adversely affect the acceptability of a targeted programme."[187] Nonetheless, the fact remains that almost 80% of respondents thought that screening was not a waste of resources, and there was only small variation between current smokers and other groups with regard to the perception that screening is a waste of money.

Even if a majority of the public regarded lung cancer screening as a waste of money, the specific ethico-legal context of Switzerland should be taken into consideration. The Swiss Constitution places great emphasis on justice and explicitly forbids discrimination on the basis of any particular condition. Article 8 of the Constitution states: "Every person is equal before the law….No person may be discriminated against, in particular on grounds of origin, race, gender, age, language, social position, way of life, religious, ideological, or political convictions, or because of a physical, mental or psychological disability." [214] Any decision not to support screening, if based even in part on the misplaced perception that lung cancer is a "self-inflicted disease", would thus be incompatible with the Swiss constitution.

However, screening for lung cancer remains unique in two ways. Even if we are careful to avoid any (mis)attribution of moral responsibility, it remains the case that many patients would not have developed lung cancer had they not smoked, and many continue to smoke despite knowing the risks, with some refusing to take part in smoking cessation programmes. These unique features may alter perceptions of overall cost-effectiveness. LDCT screening does reduce lung cancer mortality, but only 1 in 200 of those screened will avoid death from lung cancer as a result. Those who continue to smoke worsen their chances of benefitting from screening still further. At a socio-political level, perceptions of the relatively low number of lives saved by screening combined with questions about just distribution of burdens may call into question whether screening ought to be provided. Smokers and ex-smokers should not be discriminated against, but it does not follow from that conclusion that any decision not to provide screening would be discriminatory.

Finally, an important point to bear in mind is that, if LDCT screening is implemented despite the evidence suggesting that it does not reduce overall mortality and has only a small effect on lung cancer mortality, then this implementation might itself increase stigmatisation of smokers not because they are to blame for the disease (though this perception would again complicate matters), but because they could be perceived as having scarce resources 'wasted' on them for little or no benefit. It would be unfortunate if, in addition to the existing stigma that they experience, smokers and ex-smokers were subjected to further stigmatisation because of public perceptions of lung cancer screening as being ineffective. Of course, this effect would only occur if the public believed that screening is not beneficial; currently, that is not the case. However, any implementation of screening should involve disclosure to the public of what the evidence is for that decision.

19.5.3. Justice and disparities

The charge of discrimination against smokers gains more force when we consider racial and socioeconomic disparities. People from ethnic minorities and those from lower socioeconomic groups are more likely to smoke, less likely to quit, and more likely to develop lung cancer than those whiter, richer, populations [190] [193] and are also less likely to participate in screening. Furthermore, given the fact that lung cancer is more likely to affect certain populations, failing to provide a lung cancer screening programme could be seen as discriminatory given that screening exists for conditions more likely to affect wealthier and 'whiter' groups – all the more so in view of the specific Swiss legal context, as discussed in section 19.5.2.

It is certainly true that a well-implemented community-based screening programme that is able to reach those in these groups would save more lives among these specific sub-populations. Nonetheless, just as a decision not to provide screening in general could on balance not be seen as discriminatory against smokers, it seems likely that, while such a decision might be more unjust from the perspective of these groups, that unfairness would not actually amount to discrimination. If a decision regarding whether to implement a screening programme is made on the basis of cost-effectiveness (or a lack thereof) it is unlikely to be discriminatory, even if a particular sub-group of the population would benefit more from its implementation than other groups. Nonetheless, whether or not the decision about whether to provide screening is finely balanced, the beneficial effects in terms of reducing health inequalities among lower socioeconomic and minority ethnic groups should be seen as one additional advantage of screening. (Potentially, screening could be offered selectively to those in these disadvantaged ethnic and socioeconomic groups rather than to all smokers and ex-smokers – but this would in turn lead to the charge of discrimination against "better-off" smokers and ex-smokers.)

19.5.4. Justice and eligible populations

A related issue concerns the inclusion criteria for screening programmes. The eligible population is likely to include only current and former smokers, but other groups might also benefit from screening. These include those exposed to passive smoking in the workplace or at home, those with a family history of lung cancer and those exposed to other environmental factors that increase the risk of lung cancer, such as asbestos; it has been suggested that screening should also be offered to these groups. [215] Here, too, the charge of discrimination could be levelled against a screening programme that only included smokers and ex-smokers.

However, this particular objection is dismissed more straightforwardly. While a screening programme among smokers and ex-smokers might be deemed beneficial overall, the cost-effectiveness and harm-

benefit considerations are finely balanced. The large trials of LDCT screening have only included current and former smokers rather than those from other eligible populations, meaning that strong evidence exists only for these groups. Screening all those in the expanded category of passive smoking and exposure to other risk factors would substantially alter the clinical harm/benefit and economic cost-effectiveness calculations, making screening much more burdensome for only marginally increased benefit: “in screening a lower risk population, far more patients would have to be screened for every lung cancer death averted while the harms to the population would accumulate beyond the potential small overall increased benefit.” [216] In other words, the burdens imposed on these sub-populations by screening would not currently be justified by the prospective benefits of screening. Thus, clinical and cost-effectiveness considerations rather than ethical principles determine the conclusion regarding other populations than smokers and ex-smokers. However, research in this area is ongoing and screening could be expanded to other at-risk groups in the future if the evidence merits it.

19.6. Conclusion

This chapter has described and analysed the ethical issues surrounding lung cancer screening in the specific clinical context, and also from the wider perspective of justice. If screening is being implemented, patient attitudes and barriers to involving them in screening must be considered. The discussion with a potential participant in screening must adhere to best practice in terms of shared decision making and risk communication, including consideration of different screening modalities and smoking cessation and how they interact with the patient's preferences.

More broadly, despite promising results being reported in several trials, the prospective benefits of screening are sometimes overstated and inadvertently miscommunicated. This can be due to unconscious conflicts of interest. The controversial perception of smokers as being responsible for their disease is not shared by a majority of the public, and even if it were, this should not affect decision-making in this area, but it does complicate an already contentious debate. While smokers and ex-smokers should not be discriminated against, implementing screening given the relatively high costs and low benefits could actually further stigmatise this population.

Perhaps the most important point in terms of ethics is that any values must be articulated to facilitate transparency in decisions about implementing lung cancer screening at the individual patient level, the societal level and the health system level. Just as patients' values must be articulated in order to participate in shared decision making with doctors (who must themselves reveal any relevant values influencing their advice to patients), so any underlying moral values regarding justice or harm-benefit considerations must be shared so that decision-making about implementing screening is transparent and justifiable.

By their nature, decisions regarding cost-effectiveness remain contentious, and it is not clear whether screening represents fair distribution of resources given the relatively low number of lung cancer deaths prevented, the negligible effect on overall mortality and the burdens imposed on the screened population, as well as the fact that money spent on screening could be spent on other healthcare interventions and public health campaigns. As such, lung cancer screening may be neither just nor justifiable, but this judgement will depend on one's individual perspective and values as well as consideration of which other screening programmes are implemented. The individual decision about whether to participate in screening depends on the patient's personal values; similarly, the decision about whether to implement a screening programme depends on values that are not always, but should always be, articulated and discussed.

20.Overall conclusion

LDCT screening for lung cancer is associated with a reduced mortality from lung cancer but appears not to impact overall mortality. Psychological consequences of screening (e.g. anxiety or depression) remain unclear and LDCT screenings does not seem to increase quit rates from smoking. False positive findings from LDCT remain a concern and important differences in false positive rates, repeated scans and invasive work-ups were found between trials. Volumed based definitions of suspicious nodes, repeated scans and strict work-up protocols as applied in the large NELSON trial appear to reduce false positive scans.

The great majority of the published cost-effectiveness analyses concluded that lung cancer screening may be a cost-effective intervention. Analyses based on data from the NELSON trial confirmed the positive results obtained in previous analyses based on the results of the NLST. The results of the cost-effectiveness analysis suggested that most lung cancer screening strategies may be cost-effective in Switzerland (assuming a threshold of CHF 100,000 per QALY gained). The cost-effectiveness and budget impact were highly dependent on screening intervals and smoking eligibility criteria. Although being more expensive than biennial and triennial screening strategies, annual screening showed the greatest potential reduction in lung cancer mortality and the highest increase of QALY gained.

Whether lung cancer screening represents a fair distribution of harms and burdens for the benefit conferred is a subjective judgement. Even if screening is deemed cost-effective in the economic sense, there is little impact on overall mortality and the number of patients needed to screen and number of false positives incurred to prevent each lung cancer death may be too high to merit implementation. Whatever decision is ultimately made about screening, whether at the patient level or the health systems level, any values underlying that decision must be articulated clearly, along with the empirical evidence informing that decision.

21.References

1. Snowsill T, Yang H, Griffin E, et al. Low-dose computed tomography for lung cancer screening in high-risk populations: a systematic review and economic evaluation. *Health Technol Assess.* 2018;22(69):1-276.
2. Organization WH. Cancer - Key facts. *https://wwwwhoint/news-room/fact-sheets/detail/cancer.* 2019;Published September 12, 2018. Accessed December 3, 2019.
3. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer.* 2015;136(5):E359-E386.
4. Swiss Federal Statistical Office. Specific causes of death in 2016. *wwwbfsadminch/bfs/en/home/statistics/health/state-health/mortality-causes-death/specifichtml* accessed December 3, 2019.
5. NIH National Cancer Institute. Cancer Stat Facts: Lung and Bronchus Cancer. *https://seercancergov/statfacts/html/lungbhtml.*Accessed, December 3, 2019.
6. Aberle DR, Adams AM, Berg CD, et al. Reduced Lung-Cancer Mortality with Low-Dose Computed Tomographic Screening. *New England Journal of Medicine.* 2011;365(5):395-409.
7. Wille MM, Dirksen A, Ashraf H, et al. Results of the Randomized Danish Lung Cancer Screening Trial with Focus on High-Risk Profiling. *Am J Respir Crit Care Med.* 2016;193(5):542-551.
8. Huang KL, Wang SY, Lu WC, Chang YH, Su J, Lu YT. Effects of low-dose computed tomography on lung cancer screening: a systematic review, meta-analysis, and trial sequential analysis. *BMC Pulm Med.* 2019;19(1):126.
9. Yang W, Qian F, Teng J, et al. Community-based lung cancer screening with low-dose CT in China: Results of the baseline screening. *Lung Cancer.* 2018;117:20-26.
10. De Koning H, Van Der Aalst C, Ten Haaf K, Oudkerk M. PL02.05 Effects of Volume CT Lung Cancer Screening: Mortality Results of the NELSON Randomised-Controlled Population Based Trial. *Journal of Thoracic Oncology.* 2018;13(10):S185.
11. Wang X, Liu H, Shen Y, Li W, Chen Y, Wang H. Low-dose computed tomography (LDCT) versus other cancer screenings in early diagnosis of lung cancer: A meta-analysis. *Medicine (Baltimore).* 2018;97(27):e11233.
12. Coureau G, Delva F. [Lung cancer screening among the smoker population]. *Bull Cancer.* 2019;106(7-8):693-702.
13. Fu C, Liu Z, Zhu F, Li S, Jiang L. A meta-analysis: is low-dose computed tomography a superior method for risky lung cancers screening population? *The Clinical Respiratory Journal.* 2016;10(3):333-341.
14. Coureau G, Salmi LR, Etard C, Sancho-Garnier H, Sauvaget C, Mathoulin-Pélissier S. Low-dose computed tomography screening for lung cancer in populations highly exposed to tobacco: A systematic methodological appraisal of published randomised controlled trials. *European Journal of Cancer.* 2016;61:146-156.
15. Mazzone PJ, Silvestri GA, Patel S, et al. Screening for Lung Cancer: CHEST Guideline and Expert Panel Report. *Chest.* 2018;153(4):954-985.
16. Usman Ali M, Miller J, Peirson L, et al. Screening for lung cancer: A systematic review and meta-analysis. *Preventive Medicine.* 2016;89:301-314.
17. Seigneurin A, Field JK, Gachet A, Duffy SW. A systematic review of the characteristics associated with recall rates, detection rates and positive predictive values of computed tomography screening for lung cancer. *Annals of Oncology.* 2014;25(4):781-791.
18. Schütte S, Dietrich D, Montet X, Flahault A. Participation in lung cancer screening programs: are there gender and social differences? A systematic review. *Public Health Rev.* 2018;39:23-23.
19. Wang Z, Hu Y, Wang Y, et al. Can CT Screening Give Rise to a Beneficial Stage Shift in Lung Cancer Patients? Systematic Review and Meta-Analysis. *PLOS ONE.* 2016;11(10):e0164416.

20. Slatore CG, Baumann C, Pappas M, Humphrey LL. Smoking behaviors among patients receiving computed tomography for lung cancer screening. Systematic review in support of the U.S. preventive services task force. *Ann Am Thorac Soc.* 2014;11(4):619-627.
21. Piñeiro B, Simmons VN, Palmer AM, Correa JB, Brandon TH. Smoking cessation interventions within the context of Low-Dose Computed Tomography lung cancer screening: A systematic review. *Lung Cancer.* 2016;98:91-98.
22. Slatore CG, Sullivan DR, Pappas M, Humphrey LL. Patient-Centered Outcomes among Lung Cancer Screening Recipients with Computed Tomography: A Systematic Review. *Journal of Thoracic Oncology.* 2014;9(7):927-934.
23. Gierada DS, Garg K, Nath H, Strollo DC, Fagerstrom RM, Ford MB. CT Quality Assurance in the Lung Screening Study Component of the National Lung Screening Trial: Implications for Multicenter Imaging Trials. *American Journal of Roentgenology.* 2009;193(2):419-424.
24. Wong SS, Wilczynski NL, Haynes RB. Developing optimal search strategies for detecting clinically sound treatment studies in EMBASE. *J Med Libr Assoc.* 2006;94(1):41-47.
25. Gesundheitswesen IfQuWi. Lungenkrebsscreening mittels Niedrigdosis-Computertomografie. 2020.
26. Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions.* Vol Version 5.1.0. www.cochrane.org/training/cochrane-handbook The Cochrane Collaboration 2011
27. Guyatt GH, Oxman AD, Schünemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: A new series of articles in the Journal of Clinical Epidemiology. *Journal of Clinical Epidemiology.* 2011;64(4):380-382.
28. Guyatt G, Oxman AD, Akl E, et al. GRADE guidelines 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol.* 2010.
29. Guyatt G, Oxman AD, Kunz R, et al. GRADE guidelines: 2. Framing the question and deciding on important outcomes. *J Clin Epidemiol.* 2010.
30. Balshem H, Helfand M, Schunemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol.* 2011.
31. Guyatt G, Oxman AD, Vist G, et al. GRADE guidelines: 4. Rating the quality of evidence-study limitations (risk of bias) and publication bias. *J Clin Epidemiol.* 2011.
32. Guyatt GH, Oxman AD, Montori V, et al. GRADE guidelines: 5. Rating the quality of evidence--publication bias. *J Clin Epidemiol.* 2011;64(12):1277-1282.
33. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines 6. Rating the quality of evidence--imprecision. *J Clin Epidemiol.* 2011;64(12):1283-1293.
34. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 7. Rating the quality of evidence--inconsistency. *J Clin Epidemiol.* 2011;64(12):1294-1302.
35. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 8. Rating the quality of evidence--indirectness. *J Clin Epidemiol.* 2011;64(12):1303-1310.
36. Guyatt GH, Oxman AD, Sultan S, et al. GRADE guidelines: 9. Rating up the quality of evidence. *JClinEpidemiol.* 2011;64(12):1311-1316.
37. Brunetti M, Shemilt I, Pregno S, et al. Grade guidelines: 10. Considering resource use and rating the quality of economic evidence. *J Clin Epidemiol.* 2012.
38. Guyatt G, Oxman AD, Sultan S, et al. GRADE guidelines 11. Making an overall rating of confidence in effect estimates for a single outcome and for all outcomes. *J Clin Epidemiol.* 2012.
39. Guyatt GH, Oxman AD, Santesso N, et al. GRADE guidelines 12. Preparing Summary of Findings tables-binary outcomes. *J Clin Epidemiol.* 2012.
40. Guyatt GH, Thorlund K, Oxman AD, et al. GRADE guidelines: 13. Preparing Summary of Findings tables and evidence profiles - Continuous outcomes. *Journal of Clinical Epidemiology.* 2013;66(2):173-183.
41. Andrews J, Guyatt G, Oxman AD, et al. GRADE guidelines: 14. Going from evidence to recommendations: the significance and presentation of recommendations. *J Clin Epidemiol.* 2013;66(7):719-725.

42. Andrews JC, Schunemann HJ, Oxman AD, et al. GRADE guidelines: 15. Going from evidence to recommendation-determinants of a recommendation's direction and strength. *J Clin Epidemiol.* 2013;66(7):726-735.
43. Infante M, Cavuto S, Lutman FR, et al. A Randomized Study of Lung Cancer Screening with Spiral Computed Tomography. *American Journal of Respiratory and Critical Care Medicine.* 2009;180(5):445-453.
44. Rücker G, Schwarzer G, Carpenter JR, Schumacher M. Undue reliance on I(2) in assessing heterogeneity may mislead. *BMC Med Res Methodol.* 2008;8:79.
45. Dias S, Sutton AJ, Ades AE, Welton NJ. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. *Med Decis Making.* 2013;33(5):607-617.
46. Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ.* 2005;331(7521):897-900.
47. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med.* 2004;23(20):3105-3124.
48. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ.* 2021;372:n71.
49. Gohagan JK, Marcus PM, Fagerstrom RM, et al. Final results of the Lung Screening Study, a randomized feasibility study of spiral CT versus chest X-ray screening for lung cancer. *Lung Cancer.* 2005;47(1):9-15.
50. Blanchon T, Bréchot J-M, Grenier PA, et al. Baseline results of the Depiscan study: A French randomized pilot trial of lung cancer screening comparing low dose CT scan (LDCT) and chest X-ray (CXR). *Lung Cancer.* 2007;58(1):50-58.
51. Pedersen JH, Ashraf H, Dirksen A, et al. The Danish Randomized Lung Cancer CT Screening Trial—Overall Design and Results of the Prevalence Round. *Journal of Thoracic Oncology.* 2009;4(5):608-614.
52. Garg K, Keith RL, Byers T, et al. Randomized Controlled Trial with Low-Dose Spiral CT for Lung Cancer Screening: Feasibility Study and Preliminary Results. *Radiology.* 2002;225(2):506-510.
53. Pegna AL, Picozzi G, Falaschi F, et al. Four-Year Results of Low-Dose CT Screening and Nodule Management in the ITALUNG Trial. *Journal of Thoracic Oncology.* 2013;8(7):866-875.
54. Spiro SG, Hackshaw A. Research in progress—LungSEARCH: a randomised controlled trial of surveillance for the early detection of lung cancer in a high-risk group. *Thorax.* 2016;71(1):91.
55. Becker N, Motsch E, Gross ML, et al. Randomized study on early detection of lung cancer with MSCT in Germany: study design and results of the first screening round. *Journal of Cancer Research and Clinical Oncology.* 2012;138(9):1475-1486.
56. Becker N, Motsch E, Trotter A, et al. Lung cancer mortality reduction by LDCT screening-Results from the randomized German LUSI trial. *Int J Cancer.* 2019.
57. Pastorino U, Rossi M, Rosato V, et al. Annual or biennial CT screening versus observation in heavy smokers: 5-year results of the MILD trial. *European Journal of Cancer Prevention.* 2012;21(3).
58. Pastorino U, Silva M, Sestini S, et al. Prolonged lung cancer screening reduced 10-year mortality in the MILD trial: new confirmation of lung cancer screening efficacy. *Ann Oncol.* 2019;30(7):1162-1169.
59. Horeweg N, Scholten ET, de Jong PA, et al. Detection of lung cancer through low-dose CT screening (NELSON): a prespecified analysis of screening test performance and interval cancers. *The Lancet Oncology.* 2014;15(12):1342-1350.
60. Horeweg N, van der Aalst CM, Vliegenthart R, et al. Volumetric computed tomography screening for lung cancer: three rounds of the NELSON trial. *European Respiratory Journal.* 2013;42(6):1659.
61. de Koning HJ, van der Aalst CM, de Jong PA, et al. Reduced Lung-Cancer Mortality with Volume CT Screening in a Randomized Trial. *The New England journal of medicine.* 2020:10.1056/NEJMoa1911793.

62. National Lung Screening Trial Research Team, Church TR, Black WC, et al. Results of Initial Low-Dose Computed Tomographic Screening for Lung Cancer. *New England Journal of Medicine.* 2013;368(21):1980-1991.
63. Field JK, Duffy SW, Baldwin DR, 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 Technology Assessment May 2016 2016.
64. Lopes Pegna A, Picozzi G, Mascalchi M, et al. Design, recruitment and baseline results of the ITALUNG trial for lung cancer screening with low-dose CT. *Lung Cancer.* 2009;64(1):34-40.
65. Heiken JP, Brink JA, Vannier MW. Spiral (helical) CT. *Radiology.* 1993;189(3):647-656.
66. Maldjian PD, Goldman AR. Reducing Radiation Dose in Body CT: A Primer on Dose Metrics and Key CT Technical Parameters. *American Journal of Roentgenology.* 2013;200(4):741-747.
67. Sagawa M, Nakayama T, Tanaka M, Sakuma T, Sobue T, Group TJS. A Randomized Controlled Trial on the Efficacy of Thoracic CT Screening for Lung Cancer in Non-smokers and Smokers of <30 Pack-years Aged 50–64 Years (JECS Study): Research Design. *Japanese Journal of Clinical Oncology.* 2012;42(12):1219-1221.
68. Sagawa M, Sugawara T, Ishibashi N, Koyanagi A, Kondo T, Tabata T. Efficacy of low-dose computed tomography screening for lung cancer: the current state of evidence of mortality reduction. *Surgery Today.* 2017;47(7):783-788.
69. Crosbie PA, Gabe R, Simmonds I, et al. Yorkshire Lung Screening Trial (YLST): protocol for a randomised controlled trial to evaluate invitation to community-based low-dose CT screening for lung cancer versus usual care in a targeted population at risk. *BMJ Open.* 2020;10(9):e037075.
70. Field JK, Vulkan D, Davies MPA, et al. Lung cancer mortality reduction by LDCT screening: UKLS randomised trial results and international meta-analysis. *The Lancet Regional Health - Europe.* 2021;10:100179.
71. Infante M, Cavuto S, Lutman FR, et al. Long-Term Follow-up Results of the DANTE Trial, a Randomized Study of Lung Cancer Screening with Spiral Computed Tomography. *American Journal of Respiratory and Critical Care Medicine.* 2015;191(10):1166-1175.
72. Paci E, Puliti D, Lopes Pegna A, et al. Mortality, survival and incidence rates in the ITALUNG randomised lung cancer screening trial. *Thorax.* 2017;72(9):825-831.
73. Callister ME, Janes SM. Defining the path: lung cancer CT screening in Europe. *Thorax.* 2017;72(9):778.
74. Becker N, Motsch E, Gross ML, et al. Randomized Study on Early Detection of Lung Cancer with MSCT in Germany: Results of the First 3 Years of Follow-up After Randomization. *J Thorac Oncol.* 2015;10(6):890-896.
75. Field JK, Duffy SW, Baldwin DR, 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.* 2016;71(2):161.
76. Rasmussen JF, Siersma V, Pedersen JH, Brodersen J. Psychosocial consequences in the Danish randomised controlled lung cancer screening trial (DLCST). *Lung Cancer.* 2015;87(1):65-72.
77. van den Bergh KAM, Essink-Bot ML, Borsboom GJJM, Scholten ET, van Klaveren RJ, de Koning HJ. Long-term effects of lung cancer computed tomography screening on health-related quality of life: the NELSON trial. *European Respiratory Journal.* 2011;38(1):154.
78. Field JK, Duffy SW, Baldwin DR, 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 technology assessment (Winchester, England).* 2016;20:1-146.
79. Gareen IF, Duan F, Greco EM, et al. Impact of lung cancer screening results on participant health-related quality of life and state anxiety in the National Lung Screening Trial. *Cancer.* 2014;120(21):3401-3409.
80. Higgins JPT, Altman DG, Gøtzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ.* 2011;343:d5928.

81. Ashraf H, Saghir Z, Dirksen A, 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.* 2014;69(6):574.
82. Lococo F, Cardillo G, Veronesi G. Does a lung cancer screening programme promote smoking cessation? *Thorax.* 2017;72(10):870.
83. van den Bergh KAM, Essink-Bot ML, Borsboom GJJM, et al. Short-term health-related quality of life consequences in a lung cancer CT screening trial (NELSON). *British Journal of Cancer.* 2010;102(1):27-34.
84. Aggestrup LM, Hestbech MS, Siersma V, Pedersen JH, Brodersen J. Psychosocial consequences of allocation to lung cancer screening: a randomised controlled trial. *BMJ Open.* 2012;2(2):e000663.
85. Brain K, Lifford KJ, Carter B, et al. Long-term psychosocial outcomes of low-dose CT screening: results of the UK Lung Cancer Screening randomised controlled trial. *Thorax.* 2016;71(11):996.
86. Clark MA, Gorelick JJ, Sicks JD, et al. The Relations Between False Positive and Negative Screens and Smoking Cessation and Relapse in the National Lung Screening Trial: Implications for Public Health. *Nicotine & Tobacco Research.* 2016;18(1):17-24.
87. van der Aalst CM, van Klaveren RJ, van den Bergh KAM, Willemsen MC, de Koning HJ. The impact of a lung cancer computed tomography screening result on smoking abstinence. *European Respiratory Journal.* 2011;37(6):1466.
88. Brain K, Carter B, Lifford KJ, et al. Impact of low-dose CT screening on smoking cessation among high-risk participants in the UK Lung Cancer Screening Trial. *Thorax.* 2017;72(10):912.
89. Ashraf H, Tønnesen P, Holst Pedersen J, Dirksen A, Thorsen H, Døssing M. Effect of CT screening on smoking habits at 1-year follow-up in the Danish Lung Cancer Screening Trial (DLCST). *Thorax.* 2009;64(5):388.
90. Sverzellati N, Silva M, Calareso G, et al. Low-dose computed tomography for lung cancer screening: comparison of performance between annual and biennial screen. *European Radiology.* 2016;26(11):3821-3829.
91. Aberle DR, DeMello S, Berg CD, et al. Results of the Two Incidence Screenings in the National Lung Screening Trial. *New England Journal of Medicine.* 2013;369(10):920-931.
92. van Klaveren RJ, Oudkerk M, Prokop M, et al. Management of Lung Nodules Detected by Volume CT Scanning. *New England Journal of Medicine.* 2009;361(23):2221-2229.
93. Patz EF, Jr., Pinsky P, Gatsonis C, et al. Overdiagnosis in Low-Dose Computed Tomography Screening for Lung Cancer. *JAMA Internal Medicine.* 2014;174(2):269-274.
94. González Maldonado S, Motsch E, Trotter A, et al. Overdiagnosis in lung cancer screening: Estimates from the German Lung Cancer Screening Intervention Trial. *Int J Cancer.* 2021;148(5):1097-1105.
95. National Lung Screening Trial Research T. Lung Cancer Incidence and Mortality with Extended Follow-up in the National Lung Screening Trial. *J Thorac Oncol.* 2019;14(10):1732-1742.
96. Saghir Z, Dirksen A, Ashraf H, 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.* 2012;67(4):296.
97. Rasmussen JF, Siersma V, Pedersen JH, Heleno B, Saghir Z, Brodersen J. Healthcare costs in the Danish randomised controlled lung cancer CT-screening trial: A registry study. *Lung Cancer.* 2014;83(3):347-355.
98. Kubík A, Haerting J. Survival and mortality in a randomized study of lung cancer detection. *Neoplasm.* 1990;37:467–75.
99. Marcus PM, Bergstralh EJ, Fagerstrom RM, et al. Lung Cancer Mortality in the Mayo Lung Project: Impact of Extended Follow-up. *JNCI: Journal of the National Cancer Institute.* 2000;92(16):1308-1316.
100. Oken MM, Hocking WG, Kvale PA, et al. Screening by Chest Radiograph and Lung Cancer Mortality: The Prostate, Lung, Colorectal, and Ovarian (PLCO) Randomized Trial. *JAMA.* 2011;306(17):1865-1873.

101. Heijnsdijk EAM, Csanadi M, Gini A, et al. All-cause mortality versus cancer-specific mortality as outcome in cancer screening trials: A review and modeling study. *Cancer Med.* 2019;8(13):6127-6138.
102. Passiglia F, Cinquini M, Bertolaccini L, et al. Benefits and Harms of Lung Cancer Screening by Chest Computed Tomography: A Systematic Review and Meta-Analysis. *Journal of Clinical Oncology.* 2021;39(23):2574-2585.
103. Force USPST. Screening for Lung Cancer: US Preventive Services Task Force Recommendation Statement. *JAMA.* 2021;325(10):962-970.
104. Tanner NT, Dai L, Bade BC, Gebregziabher M, Silvestri GA. Assessing the Generalizability of the National Lung Screening Trial: Comparison of Patients with Stage 1 Disease. *Am J Respir Crit Care Med.* 2017;196(5):602-608.
105. Evers S, Goossens M, de Vet H, van Tulder M, Ament A. Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria. . *Int J Technol Assess Health Care* 2005;21:240–5.
106. Drummond M, Barbieri M, Cook J, et al. Transferability of economic evaluations across jurisdictions: ISPOR Good Research Practices Task Force report. *Value Health.* 2009;12:409-418.
107. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the economic evaluation of health care programme. Third edition. 2005.
108. O'Brien BJ. A tale of two (or more) cities: geographic transferability of pharmacoeconomic data. *Am J Manag Care.* 1997;3 Suppl:S33-39.
109. Welte R, Feenstra T, Jager H, Leidl R. A decision chart for assessing and improving the transferability of economic evaluation results between countries. *Pharmacoeconomics.* 2004;22:857-876.
110. Black WC. Computed Tomography Screening for Lung Cancer in the National Lung Screening Trial: A Cost-effectiveness Analysis. *Journal of Thoracic Imaging.* 2015;30(2).
111. Black WC, Gareen IF, Soneji SS, et al. Cost-Effectiveness of CT Screening in the National Lung Screening Trial. *New England Journal of Medicine.* 2014;371(19):1793-1802.
112. Chirikos TN, Hazelton T, Tockman M, Clark R. Screening for Lung Cancer With CT: A Preliminary Cost-effectiveness Analysis. *Chest.* 2002;121(5):1507-1514.
113. Goffin JR, Flanagan WM, Miller AB, et al. Cost-effectiveness of Lung Cancer Screening in Canada. *JAMA Oncology.* 2015;1(6):807-813.
114. Goffin JR, Flanagan WM, Miller AB, et al. Biennial lung cancer screening in Canada with smoking cessation-outcomes and cost-effectiveness. *Lung cancer (Amsterdam, Netherlands).* 2016;101:98-103.
115. Goulart BHL, Bensink ME, Mummy DG, Ramsey SD. Lung cancer screening with low-dose computed tomography: costs, national expenditures, and cost-effectiveness. *Journal of the National Comprehensive Cancer Network : JNCCN.* 2012;10:267-275.
116. Mahadevia PJ, Fleisher LA, Frick KD, Eng J, Goodman SN, Powe NR. Lung Cancer Screening With Helical Computed Tomography in Older Adult Smokers: A Decision and Cost-effectiveness Analysis. *JAMA.* 2003;289(3):313-322.
117. Manser R, Dalton A, Carter R, Byrnes G, Elwood M, Campbell DA. Cost-effectiveness analysis of screening for lung cancer with low dose spiral CT (computed tomography) in the Australian setting. *Lung Cancer.* 2005;48(2):171-185.
118. Marshall D, Simpson KN, Earle CC, Chu C-W. Potential cost-effectiveness of one-time screening for lung cancer (LC) in a high risk cohort. *Lung Cancer.* 2001;32(3):227-236.
119. McMahon PM, Kong CY, Bouzan C, et al. Cost-Effectiveness of Computed Tomography Screening for Lung Cancer in the United States. *Journal of Thoracic Oncology.* 2011;6(11):1841-1848.
120. Pyenson BS, Henschke CI, Yankelevitz DF, Yip R, Dec E. Offering lung cancer screening to high-risk medicare beneficiaries saves lives and is cost-effective: an actuarial analysis. *American health & drug benefits.* 2014;7:272-282.

121. Pyenson BS, Sander MS, Jiang Y, Kahn H, Mulshine JL. An Actuarial Analysis Shows That Offering Lung Cancer Screening As An Insurance Benefit Would Save Lives At Relatively Low Cost. *Health Affairs.* 2012;31(4):770-779.
122. Report C. Low-Dose Computed Tomography for Lung Cancer Screening: A Review of the Clinical Effectiveness, Diagnostic Accuracy, Cost-Effectiveness, and Guidelines. 2015.
123. Shmueli A, Fraifeld S, Peretz T, et al. Cost-Effectiveness of Baseline Low-Dose Computed Tomography Screening for Lung Cancer: The Israeli Experience. *Value in Health.* 2013;16(6):922-931.
124. Tabata H, Akita T, Matsuura A, et al. Cost-effectiveness of the introduction of low-dose CT screening in Japanese smokers aged 55 to 74 years old. *Hiroshima J Med Sci.* 2014;63:13–22.
125. ten Haaf K, Tammemägi MC, Bondy SJ, 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 Medicine.* 2017;14(2):e1002225.
126. Villanti AC, Jiang Y, Abrams DB, Pyenson BS. A Cost-Utility Analysis of Lung Cancer Screening and the Additional Benefits of Incorporating Smoking Cessation Interventions. *PLOS ONE.* 2013;8(8):e71379.
127. Whynes DK. Could CT screening for lung cancer ever be cost effective in the United Kingdom? *Cost Effectiveness and Resource Allocation.* 2008;6(1):5.
128. Wisnivesky JP, Mushlin AI, Sicherman N, Henschke C. The Cost-Effectiveness of Low-Dose CT Screening for Lung Cancer*: Preliminary Results of Baseline Screening. *Chest.* 2003;124(2):614-621.
129. Cressman S, Peacock SJ, Tammemagi MC, et al. The Cost-Effectiveness of High-Risk Lung Cancer Screening and Drivers of Program Efficiency. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer.* 2017;12:1210-1222.
130. Evans WK, Flanagan WM, Miller AB, et al. Implementing low-dose computed tomography screening for lung cancer in Canada: implications of alternative at-risk populations, screening frequency, and duration. *Current oncology (Toronto, Ont).* 2016;23(3):e179-187.
131. Treskova M, Aumann I, Golpon H, Vogel-Claussen J, Welte T, Kuhlmann A. Trade-off between benefits, harms and economic efficiency of low-dose CT lung cancer screening: a microsimulation analysis of nodule management strategies in a population-based setting. *BMC medicine.* 2017;15:162.
132. Goffin JR, Flanagan WM, Miller AB, et al. Biennial lung cancer screening in Canada with smoking cessation—outcomes and cost-effectiveness. *Lung Cancer.* 2016;101:98-103.
133. Diaz M, Garcia M, Vidal C, et al. Health and economic impact at a population level of both primary and secondary preventive lung cancer interventions: A model-based cost-effectiveness analysis. *Lung Cancer.* 2021;159:153-161.
134. Jaine R, Kvizhinadze G, Nair N, Blakely T. Cost-effectiveness of a low-dose computed tomography screening programme for lung cancer in New Zealand. *Lung cancer (Amsterdam, Netherlands).* 2018;124:233-240.
135. Kim J, Cho B, Kim S-H, Choi C-M, Kim Y, Jo M-W. Cost Utility Analysis of a Pilot Study for the Korean Lung Cancer Screening Project. *Cancer Research and Treatment.* 2021:1-29.
136. Kumar V, Cohen JT, van Klaveren D, et al. Risk-Targeted Lung Cancer Screening: A Cost-Effectiveness Analysis. *Annals of internal medicine.* 2018;168:161-169.
137. Wade S, Weber M, Caruana M, et al. Estimating the Cost-Effectiveness of Lung Cancer Screening with Low-Dose Computed Tomography for High-Risk Smokers in Australia. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer.* 2018;13:1094-1105.
138. Yang S-C, Lai W-W, Lin C-C, et al. Cost-effectiveness of implementing computed tomography screening for lung cancer in Taiwan. *Lung cancer (Amsterdam, Netherlands).* 2017;108:183-191.
139. Pyenson BS, Henschke CI, Yankelevitz DF, 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. 2014;7:272–82.
140. Tabata H, Akita T, Matsuura A, et al. Cost-effectiveness of the introduction of low-dose CT screening in Japanese smokers aged 55 to 74 years old. *Hiroshima journal of medical sciences.* 2014;63:13-22.
141. Yuan J, Sun Y, Wang K, et al. Cost Effectiveness of Lung Cancer Screening With Low-Dose CT in Heavy Smokers in China. *Cancer Prevention Research.* 2021.
142. Esmaeili MH, Seyednejad F, Mahboub-Ahari A, et al. Cost-effectiveness analysis of lung cancer screening with low-dose computed tomography in an Iranian high-risk population. *J Med Screen.* 2021;28(4):494-501.
143. Sun C, Zhang X, Guo S, et al. Determining cost-effectiveness of lung cancer screening in urban Chinese populations using a state-transition Markov model. *BMJ Open.* 2021;11(7):1-9.
144. Veronesi G, Navone N, Novellis P, et al. Favorable incremental cost-effectiveness ratio for lung cancer screening in Italy. *Lung cancer (Amsterdam, Netherlands).* 2020;143:73-79.
145. Institute of Health Economics. Low dose computed tomography for the screening of lung cancer in adults. Edmonton AB: Institute of Health Economics. 2014.
146. Criss SD, Cao P, Bastani M, et al. Cost-Effectiveness Analysis of Lung Cancer Screening in the United States: A Comparative Modeling Study. *Annals of internal medicine.* 2019.
147. Hofer F, Kauczor H-U, Stargardt T. Cost-utility analysis of a potential lung cancer screening program for a high-risk population in Germany: A modelling approach. *Lung cancer (Amsterdam, Netherlands).* 2018;124:189-198.
148. Tomonaga Y, Ten Haaf K, Frauenfelder T, 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 (Amsterdam, Netherlands).* 2018;121:61-69.
149. Du Y, Sidorenkov G, Heuvelmans MA, et al. Cost-effectiveness of lung cancer screening with low-dose computed tomography in heavy smokers: a microsimulation modelling study. *Eur J Cancer.* 2020;135:121-129.
150. Hinde S, Crilly T, Balata H, et al. The cost-effectiveness of the Manchester 'lung health checks', a community-based lung cancer low-dose CT screening pilot. *Lung cancer (Amsterdam, Netherlands).* 2018;126:119-124.
151. McLeod M, Sandiford P, Kvizhinadze G, Bartholomew K, Crengle S. Impact of low-dose CT screening for lung cancer on ethnic health inequities in New Zealand: a cost-effectiveness analysis. *BMJ open.* 2020;10(9):e037145-e037145.
152. Toumazis I, Tsai EB, Erdogan SA, et al. Cost-Effectiveness Analysis of Lung Cancer Screening Accounting for the Effect of Indeterminate Findings. *JNCI cancer spectrum.* 2019;3(3):pkz035-pkz035.
153. Gómez-Carballo N, Fernández-Soberón S, Rejas-Gutiérrez J. Cost-effectiveness analysis of a lung cancer screening programme in Spain. *European Journal of Cancer Prevention.* 2021;Publish Ahead of Print(Lc):1-10.
154. Griffin E, Hyde C, Long L, et al. Lung cancer screening by low-dose computed tomography: a cost-effectiveness analysis of alternative programmes in the UK using a newly developed natural history-based economic model. *Diagn Progn Res.* 2020;4(1):20-20.
155. Tammemägi MC, Ruparel M, Tremblay A, et al. USPSTF2013 versus PLCOm2012 lung cancer screening eligibility criteria (International Lung Screening Trial): interim analysis of a prospective cohort study. *The Lancet Oncology.* 2021.
156. ten Haaf K, van Rosmalen J, de Koning HJ. Lung Cancer Detectability by Test, Histology, Stage, and Gender: Estimates from the NLST and the PLCO Trials. *Cancer Epidemiology Biomarkers & Prevention.* 2015;24(1):154.
157. Meza R, ten Haaf K, Kong CY, et al. Comparative analysis of 5 lung cancer natural history and screening models that reproduce outcomes of the NLST and PLCO trials. *Cancer.* 2014;120(11):1713-1724.
158. Heidenreich WF, Luebeck EG, Moolgavkar SH. Some properties of the hazard function of the two-mutation clonal expansion model. *Risk Anal.* 1997;17(3):391-399.

159. Meza R, Hazelton WD, Colditz GA, Moolgavkar SH. Analysis of lung cancer incidence in the nurses' health and the health professionals' follow-up studies using a multistage carcinogenesis model. *Cancer Causes & Control.* 2008;19(3):317-328.
160. Yau G, Lock M, Rodrigues G. Systematic review of baseline low-dose CT lung cancer screening. *Lung Cancer.* 2007;58(2):161-170.
161. Office FS. - Ständige Wohnbevölkerung nach Alter, Geschlecht und Staatsangehörigkeitskategorie, 2010-2020. In:Accessed 24 February 2022.
162. BFS, Schweizerische Gesundheitsbefragung 1992, 1997, 2002, 2007, and 2012.
163. Yao YW, Yuan DM, Lu YL, Li YF, Song Y. [Screening of early lung cancer with low-dose computed tomography in high-risk populations: a meta-analysis]. *Zhonghua Yi Xue Za Zhi.* 2011;91(40):2819-2823.
164. Holford TR, Levy DT, McKay LA, et al. Patterns of birth cohort-specific smoking histories, 1965-2009. *Am J Prev Med.* 2014;46(2):e31-e37.
165. Levy DT, Romano E, Mumford E. The Relationship of Smoking Cessation to Sociodemographic Characteristics, Smoking Intensity, and Tobacco Control Policies. *Nicotine & Tobacco Research.* 2005;7(3):387-396.
166. Fiore MC, Novotny TE, Pierce JP, et al. Methods used to quit smoking in the United States. Do cessation programs help? *Jama.* 1990;263(20):2760-2765.
167. Anderson CM, Burns DM, Dodd KW, Feuer EJ. Chapter 2: Birth-cohort-specific estimates of smoking behaviors for the U.S. population. *Risk Anal.* 2012;32 Suppl 1(0 1):S14-S24.
168. Warner KE. Possible Increases in the Underreporting of Cigarette Consumption. *Journal of the American Statistical Association.* 1978;73(362):314-318.
169. Hatziandreu EJ, Pierce JP, Fiore MC, Grise V, Novotny TE, Davis RM. The reliability of self-reported cigarette consumption in the United States. *Am J Public Health.* 1989;79(8):1020-1023.
170. Jakob J, Cornuz J, Diethelm P. Prevalence of tobacco smoking in Switzerland: do reported numbers underestimate reality? *Swiss Med Wkly.* 2017;147:w14437.
171. Rosenberg MA, Feuer EJ, Yu B, et al. Chapter 3: Cohort life tables by smoking status, removing lung cancer as a cause of death. *Risk Anal.* 2012;32 Suppl 1(Suppl 1):S25-38.
172. Thun MJ, Carter BD, Feskanich D, et al. 50-year trends in smoking-related mortality in the United States. *N Engl J Med.* 2013;368(4):351-364.
173. Tomonaga Y, ten Haaf K, Frauenfelder T, 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.* 2018;121:61-69.
174. ten Haaf K, J J, MC T. Risk prediction models for selection of lung cancer screening candidates: A retrospective validation study. *PLOS Medicine.* 2017;14(4):e1002277.
175. Blom EF, Haaf Kt, de Koning HJ. Systematic Review and Meta-Analysis of Community- and Choice-Based Health State Utility Values for Lung Cancer. *PharmacoEconomics.* 2020;38(11):1187-1200.
176. Huber MB, Felix J, Vogelmann M, Leidl R. Health-Related Quality of Life of the General German Population in 2015: Results from the EQ-5D-5L. *International journal of environmental research and public health.* 2017;14(4).
177. Janssen MF, Szende A, Cabases J, Ramos-Goñi JM, Vilagut G, König HH. Population norms for the EQ-5D-3L: a cross-country analysis of population surveys for 20 countries. *The European journal of health economics : HEPAC : health economics in prevention and care.* 2019;20(2):205-216.
178. Luthy C, Cedraschi C, Allaz AF, Herrmann FR, Ludwig C. Health status and quality of life: results from a national survey in a community-dwelling sample of elderly people. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation.* 2015;24(7):1687-1696.

179. N'Goran AA, Déruaz-Luyet A, Haller DM, et al. Comparing the self-perceived quality of life of multimorbid patients and the general population using the EQ-5D-3L. *PLoS One.* 2017;12(12):e0188499.
180. Perneger TV, Combescure C, Courvoisier DS. General population reference values for the French version of the EuroQol EQ-5D health utility instrument. *Value Health.* 2010;13(5):631-635.
181. Meza R, Jeon J, Toumazis I, et al. Evaluation of the Benefits and Harms of Lung Cancer Screening With Low-Dose Computed Tomography: Modeling Study for the US Preventive Services Task Force. *JAMA.* 2021;325(10):988-997.
182. Ten Haaf K, Bastani M, Cao P, et al. A Comparative Modeling Analysis of Risk-Based Lung Cancer Screening Strategies. *J Natl Cancer Inst.* 2020;112(5):466-479.
183. Baldwin DR, Brain K, Quaife S. Participation in lung cancer screening. *Transl Lung Cancer Res.* 2021;10(2):1091-1098.
184. Bulliard J-L BK, Fracheboud J, Zwahlen M. . *Breast cancer screening programmes in Switzerland, 2010-2018, Final Report.* February 2021.
185. Richard A, Rohrmann S, Schmid SM, et al. Lifestyle and health-related predictors of cervical cancer screening attendance in a Swiss population-based study. *Cancer Epidemiol.* 2015;39(6):870-876.
186. Woloshin S, Schwartz LM, Black WC, Kramer BS. Cancer screening campaigns--getting past uninformative persuasion. *N Engl J Med.* 2012;367(18):1677-1679.
187. Quaife SL, Vrinten C, Ruparel M, et al. Smokers' interest in a lung cancer screening programme: a national survey in England. *BMC Cancer.* 2018;18(1):497.
188. Smits SE, McCutchan GM, Hanson JA, Brain KE. Attitudes towards lung cancer screening in a population sample. *Health Expect.* 2018;21(6):1150-1158.
189. Kellen E, Gabriels S, Van Hal G, Goossens MC. Lung cancer screening: intention to participate and acceptability among Belgian smokers. *Eur J Cancer Prev.* 2021;30(6):457-461.
190. Quaife SL, Marlow LAV, McEwen A, Janes SM, Wardle J. Attitudes towards lung cancer screening in socioeconomically deprived and heavy smoking communities: informing screening communication. *Health Expect.* 2017;20(4):563-573.
191. Silvestri GA, Nietert PJ, Zoller J, Carter C, Bradford D. Attitudes towards screening for lung cancer among smokers and their non-smoking counterparts. *Thorax.* 2007;62(2):126-130.
192. Chambers SK, Dunn J, Occhipinti S, et al. A systematic review of the impact of stigma and nihilism on lung cancer outcomes. *BMC Cancer.* 2012;12:184.
193. Raman V, Yong V, Erkmen CP, Tong BC. Social Disparities in Lung Cancer Risk and Screening. *Thorac Surg Clin.* 2022;32(1):23-31.
194. Beauchamp TL, Childress JF. *Principles of biomedical ethics.* 6 ed: New York: Oxford University Press; 2009.
195. (NICE) NIfHaCE. Shared decision making - NICE guideline [NG197]. Published 2021. Accessed 4.2.2022, 2022.
196. Shaw D, Elger B. Evidence-based persuasion: an ethical imperative. *Jama.* 2013;309(16):1689-1690.
197. Pinsky PF. Assessing the benefits and harms of low-dose computed tomography screening for lung cancer. *Lung Cancer Manag.* 2014;3(6):491-498.
198. Rasmussen JF, Siersma V, Malmqvist J, Brodersen J. Psychosocial consequences of false positives in the Danish Lung Cancer CT Screening Trial: a nested matched cohort study. *BMJ Open.* 2020;10(6):e034682.
199. Balata H, Tonge J, Barber PV, et al. Attendees of Manchester's Lung Health Check pilot express a preference for community-based lung cancer screening. *Thorax.* 2019;74(12):1176-1178.
200. Barta JA, Shusted CS, Ruane B, et al. Racial Differences in Lung Cancer Screening Beliefs and Screening Adherence. *Clin Lung Cancer.* 2021;22(6):570-578.
201. Heuvelmans MA, Oudkerk M. Appropriate screening intervals in low-dose CT lung cancer screening. *Transl Lung Cancer Res.* 2018;7(3):281-287.

202. Bartlett EC, Belsey J, Derbyshire J, et al. Implications of incidental findings from lung screening for primary care: data from a UK pilot. *npj Primary Care Respiratory Medicine.* 2021;31(1):36.
203. Parker L, Carter S, Williams J, Pickles K, Barratt A. Avoiding harm and supporting autonomy are under-prioritised in cancer-screening policies and practices. *Eur J Cancer.* 2017;85:1-5.
204. Mahadevia PJ, Fleisher LA, Frick KD, Eng J, Goodman SN, Powe NR. Lung cancer screening with helical computed tomography in older adult smokers: a decision and cost-effectiveness analysis. *Jama.* 2003;289(3):313-322.
205. Kyle C. Screening for lung cancer: too much for too little? *Virtual Mentor.* 2006;8(1):30-33.
206. Jonas DE, Reuland DS, Reddy SM, et al. Screening for Lung Cancer With Low-Dose Computed Tomography: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. *Jama.* 2021;325(10):971-987.
207. Biller-Andorno N, Jüni P. Abolishing mammography screening programs? A view from the Swiss Medical Board. *N Engl J Med.* 2014;370(21):1965-1967.
208. Kinsinger LS, Anderson C, Kim J, et al. Implementation of Lung Cancer Screening in the Veterans Health Administration. *JAMA Intern Med.* 2017;177(3):399-406.
209. Redberg RF, O'Malley PG. Important Questions About Lung Cancer Screening Programs When Incidental Findings Exceed Lung Cancer Nodules by 40 to 1. *JAMA Intern Med.* 2017;177(3):311-312.
210. Neeser K, Szucs T, Bulliard JL, Bachmann G, Schramm W. Cost-effectiveness analysis of a quality-controlled mammography screening program from the Swiss statutory health-care perspective: quantitative assessment of the most influential factors. *Value Health.* 2007;10(1):42-53.
211. Raftery J, Williams HC, Clarke A, et al. 'Not clinically effective but cost-effective' - paradoxical conclusions in randomised controlled trials with 'doubly null' results: a cross-sectional study. *BMJ Open.* 2020;10(1):e029596.
212. Marseille E, Kahn JG. Utilitarianism and the ethical foundations of cost-effectiveness analysis in resource allocation for global health. *Philosophy, Ethics, and Humanities in Medicine.* 2019;14(1):5.
213. Bilinski A, Neumann P, Cohen J, Thorat T, McDaniel K, Salomon JA. When cost-effective interventions are unaffordable: Integrating cost-effectiveness and budget impact in priority setting for global health programs. *PLOS Medicine.* 2017;14(10):e1002397.
214. Confederation TS. Federal Constitution of the Swiss Confederation; Article 8.2 Published 1999. Accessed 4.2.2022, 2022.
215. Wood DE. POINT: Should Lung Cancer Screening Be Expanded to Persons Who Don't Currently Meet Accepted Criteria Set Forth by the CHEST Guidelines on Lung Cancer Screening? Yes. *Chest.* 2018;153(6):1299-1302.
216. Mazzone PJ. COUNTERPOINT: Should Lung Cancer Screening Be Expanded to Persons Who Don't Currently Meet Accepted Criteria Set Forth by the CHEST Guidelines on Lung Cancer Screening? No. *Chest.* 2018;153(6):1303-1305.
217. Wille MMW, Dirksen A, Ashraf H, et al. Results of the Randomized Danish Lung Cancer Screening Trial with Focus on High-Risk Profiling. *American Journal of Respiratory and Critical Care Medicine.* 2015;193(5):542-551.
218. Kubík A, Polák J. Lung cancer detection. Results of a randomized prospective study in Czechoslovakia. *Cancer.* 1986;57(12):2427-2437.
219. Fontana RS, Sanderson DR, Woolner LB, et al. Screening for lung cancer. A critique of the mayo lung project. *Cancer.* 1991;67(S4):1155-1164.
220. Marcus PM, Prorok PC. Reanalysis of the Mayo Lung Project data: the impact of confounding and effect modification. *J Med Screen.* 1999;6(1):47-49.

22.Appendix 1 – Clinical effectiveness literature search strategy

Overview	
Internal UBM ID	DOKU_Suche-Bucher_HTA-Lung-Cancer_20201203_hae
Researcher	
Study title (preliminary)	Low-dose computed tomography for lung cancer screening in high-risk populations: a systematic review and economic evaluation Update of the 2018 HTA: *Snowsill T, Yang H, Griffin E, Long L, Varley-Campbell J, Coelho H, et al. Low-dose computed tomography for lung cancer screening in high-risk populations: a systematic review and economic evaluation. Health Technol Assess 2018;22(69).*
Research aim	To assess the clinical effectiveness and cost-effectiveness of lung cancer screening in a high-risk population using LDCT screening
The following is copied from the HTA report:	
Population	People identified as being at 'high' risk of **lung cancer**. The eligible population was individuals at high risk of lung cancer. Any definitions of high-risk populations were eligible in order to facilitate exploration of risk as a particular feature by which clinical effectiveness and cost-effectiveness might vary.
Intervention	**Low-dose CT screening**. Low-dose CT screening programmes, including both single and multiple rounds, were eligible for inclusion. We carefully investigated variations in the screening programme, not only in the techniques used to do the initial screen, but also the criteria used to define positive tests and how positive and indeterminate tests (when applicable) were followed up.
Comparison	No screening was set by the scope as the primary comparator. We have also included alternative screening programmes (e.g. CXR) for comparative purposes. The eligible comparators were **usual care** (no screening) or **other imaging** technology screening programmes (such as CXR), including both single and multiple screening rounds.
Outcome	From the scope, the outcomes suggested were potential effect on mortality, QoL and cost-effectiveness. Additional outcomes that were deemed relevant following consultation with our advisory committee included lung cancer incidence, stage and morphology of lung cancer, follow-up investigations and treatments, smoking cessation, adherence to screening, diagnostic accuracy, radiation dose of screening and adverse psychological impacts. The following outcomes were included: – lung cancer mortality – all-cause mortality – stage distributions of lung cancers – number of lung cancers detected

	– number and type of follow-up investigations – number of patients who were more amenable to surgicatreatment – surgicaresection rate – any HRQoL – smoking cessation and patients' smoking behaviour change – adherence rate to screening – diagnostic accuracy outcomes (including indeterminate results) – overdiagnosis – complications in those who underwent an invasive procedure – radiation dose of screening – radiation-related patient outcomes – adverse psychologicaimpact.
Study design	The eligible study design was **RCTs**. The following types of report were excluded: editorials and opinions, case reports and reports focusing on only technical aspects of the CT technology (such as technical descriptions of the CT technology).
Exclusion criteria	– animal models – preclinical and biological studies – non-systematic reviews, editorials, opinions – non-English language papers – reports published as meeting abstracts only, as there is unlikely to be sufficient methodological details – to allow critical appraisal of study quality.
Databases	Medline, Embase, Web of Science, Cochrane, CINAHL
Other	Search for ongoing clinical trials: in 2 Registries (clinicaltrials.gov, the WHO registry) done by HCB

Table A1 Search strategy based on PICO question, study design inclusion and exclusion criteria

Medline via Ovid (Ovid MEDLINE(R) ALL), date of search: 17.12.2020; date parameters: 1946 to December 15, 2020; 3079 hits

1. exp Lung Neoplasms/

2. ((lung$ or bronch$ or pulmon$) adj3 (cancer$ or neopla$ or tumor$ or tumour$ or carcinoma$ or adenocarcinoma$ or small cell or squamous)).ti,ab,ot,kw.

3. (NSLC or NSCLC or SLC or SCLC).ti,ab,ot,kw.

4. 1 or 2 or 3

5. exp Tomography, X-Ray Computed/

6. exp Radiography, Thoracic/

7. (x ray or xray or x-ray or CXR or radiograph$).ti,ab,ot,kw.

8. ((CT or CAT) adj3 (scan$ or screen$)).ti,ab,ot,kw.

9. ((computer$ adj3 tomogra$) and (scan$ or screen$)).ti,ab,ot,kw.

10. (tomogra$ or helix or helical or spiral$ or spiro$).ti,ab,ot,kw.

11. 5 or 6 or 7 or 8 or 9 or 10

12. 4 and 11

13. (2016* or 2017* or 2018* or 2019* or 2020* or 2021*).dt,ez,ep,up,rd,ed.

14. randomized controlled trial.pt.

15. controlled clinical trial.pt.

16. randomized.ab.

17. placebo.ab.

18. drug therapy.fs.

19. randomly.ab.

20. trial.ab.

21. groups.ab.

22. 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21

23. exp animals/ not humans.sh.

24. 22 not 23

25. 12 and 13 and 24

Note: Lines 14 to 21 are the Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity-maximizing version (2008 revision); Ovid format: Lefebvre C, Manheimer E, Glanville J. Chapter 6: Searching for studies. In: Higgins J, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org

EMBASE via Ovid, date of search: 17.12.2020; date parameters: 1974 to 2020 December 15; 3016 hits

1. exp lung cancer/ or exp lung tumor/ or exp bronchus cancer/

2. ((lung$ or bronch$ or pulmon$) adj3 (cancer$ or neopla$ or tumor$ or tumour$ or carcinoma$ or adenocarcinoma$ or small cell or squamous)).ti,ab,ot,kw.

3. (NSLC or NSCLC or SLC or SCLC).ti,ab,ot,kw.

4. 1 or 2 or 3

5. exp computer assisted tomography/

6. ((CT or CAT) adj3 (scan$ or screen$)).ti,ab,ot,kw.

7. ((computer$ adj3 tomogra$) and (scan$ or screen$)).ti,ab,ot,kw.

8. (tomogra$ or helix or helical or spiral$ or spiro$).ti,ab,ot,kw.

9. exp thorax radiography/

10. (x ray or xray or x-ray or CXR or radiograph$).ti,ab,ot,kw.

11. 5 or 6 or 7 or 8 or 9 or 10

12. 4 and 11

13. Clinical trial/

14. Randomized controlled trial/

15. Randomization/

16. Single blind procedure/

17. Double blind procedure/

18. Crossover procedure/

19. Placebo/

20. Randomi?ed controlled trial$.tw.

21. Rct.tw.

22. Random allocation.tw.

23. Randomly allocated.tw.

24. Allocated randomly.tw.

25. (allocated adj2 random).tw.

26. Single blind$.tw.

27. Double blind$.tw.

28. ((treble or triple) adj1 blind$).tw.

29. Placebo$.tw.

30. Prospective study/

31. or/13-30

32. Case study/

33. Case report.tw.

34. Abstract report/ or letter/

35. 32 or 33 or 34

36. 31 not 35

37. 12 and 36

38. (2016* or 2017* or 2018* or 2019* or 2020* or 2021*).dp,rd,dr,dc,dd.

39. 37 and 38

Note: The RCT strategy is adapted from the Scottish Intercollegiate Guidelines Network (SIGN) filter: The Scottish Intercollegiate Guidelines Network (SIGN) Search Filters Web Page. Available from: https://www.sign.ac.uk/what-we-do/methodology/search-filters/

Web of Science via Clarivate Analytics, date of search: 17.12.2020; date parameters: 2016-2020; 2931 hits

Database: Web of Science Core Collection: Science Citation Index (SCI) and Conference Proceedings Citation Index –Science (CPCI-S).

1. TS=((lung* or bronch* or pulmon*) near/3 (cancer* or neopla* or tumor* or tumour* or carcinoma* or adenocarcinoma* or "small cell" or squamous))

2. TS=(NSLC or NSCLC or SLC or SCLC)

3. TS=((CT or CAT) near/3 (scan* or screen*))

4. TS=((computer* near/3 tomogram*) and (scan* or screen*))

5. TS=(tomogra* or helix or helical or spiral* or spiro*)

6. TS=(x ray or xray or x-ray or CXR or radiograph*)

7. #1 or #2

8. #3 or #4 or #5 OR #6

9. #7 and #8

10. TS= clinical trial* OR TS=research design OR TS=comparative stud* OR TS=evaluation stud* OR TS=controlled trial* OR TS=follow-up stud* OR TS=prospective stud* OR TS=random* OR TS=placebo* OR TS=(single blind*) OR TS=(double blind*)

11. #9 and #10

Note: TS searches in the Abstract, Title, and/or Keywords fields of a record; Indexes=SCI-EXPANDED, CPCI-S Timespan=2016-2020

The Cochrane Central Register of Controlled Trials (CENTRAL) via Cochrane Collaboration; date of search: 17.12.2020; date parameters: 2016-2020; 1093 hits

#1 MeSH descriptor: [Lung Neoplasms] explode all trees

#2 ((lung* or bronch* or pulmon*) near/3 (cancer* or neopla* or tumor* or tumour* or carcinoma* or adenocarcinoma* or "small cell" or squamous)):ti,ab,kw

#3 (NSLC or NSCLC or SLC or SCLC):ti,ab,kw

#4 #1 or #2 or #3

#5 MeSH descriptor: [Tomography, X-Ray Computed] explode all trees

#6 ((CT or CAT) near/3 (scan* or screen*)):ti,ab,kw

#7 ((computer* near/3 tomogra*) and (scan* or screen*)):ti,ab,kw

#8 (tomogra* or helix or helical or spiral* or spiro*):ti,ab,kw

#9 MeSH descriptor: [Radiography, Thoracic] explode all trees

#10 (x ray or xray or x-ray or CXR or radiograph*):ti,ab,kw

#11 #5 or #6 or #7 or #8 or #9 or #10

#12 #4 and #11 Publication Year from 2016 to 2020, in Trials

Note:

CINAHL via EBSCOhost, date of search: 17.12.2020; date parameters: January 2016 – December 2020; 811 hits

1. (MH "Lung Neoplasms+")

2. TX (lung* or bronch* or pulmon*) N3 (cancer* or neopla* or tumor* or tumour* or carcinoma* or adenocarcinoma* or "small cell" or squamous)

3. TX (NSLC or NSCLC or SLC or SCLC)

4. S1 OR S2 OR S3

5. (MH "Tomography, X-Ray Computed+")

6. TX (CT or CAT) N3 (scan* or screen*)

7. TX (computer* N3 tomogra*) and (scan* or screen*)

8. TX (tomogra* or helix or helical or spiral* or spiro*)

9. TX (x ray or xray or x-ray or CXR or radiograph*)

10. (MH “Radiography, Thoracic+”)

11. S5 and S6 and S7 and S8 and S9

12. S4 AND S11

13. (MH “Clinical Trials+”)

14. PT Clinical Trial

15. TX clinic* n1 trial*

16. TX ((singl* n1 blind*) or (singl* n1 mask*))

17. TX ((doubl* n1 blind*) or (doubl* n1 mask*))

18. TX ((tripl* n1 blind*) or (tripl* n1 mask*))

19. TX ((trebl* n1 blind*) or (trebl* n1 mask*))

20. TX randomi* control* trial*

21. (MH “Random Assignment”)

22. TX random* allocat*

23. TX placebo*

24. (MH “Placebos”)

25. (MH “Quantitative Studies”)

26. TX allocat* random*

27. S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24 or S25 or S26 or S27 or S28

28. S12 AND S27

29. Limit to 2016-December 2020

Note: The RCT strategy is from the Scottish Intercollegiate Guidelines Network (SIGN) filter: The Scottish Intercollegiate Guidelines Network (SIGN) Search Filters Web Page. Available from: https://www.sign.ac.uk/what-we-do/methodology/search-filters/

23.Appendix 2 – Characteristics of the included studies

Study identifier	Arm	Country and number of centres	Number of patients approached	Number of patients randomised	Number of patients screened at baseline (*n/N*, %)	Characteristics of patients at baseline				
						Median age (years), (range)	Male, n (%) intervention/ control	Current smokers n (%) intervention/ control	Former smokers n (%) intervention/ control	Family history of lung cancer, n (%) intervention/ control
DANTE [43]	LDCT	Italy, 3 centres	2811	2811 (1403 vs. 1408)	1276	64.3 (64.0–64.7)	1276 (100) 1186 (100)	714 (56.5) 681 (57.4)	NR NR	NR NR
	Control				1196	64.6 (64.3–64.9)	1196/1196 (100)	681/1196 (56.9)	515/1196 (43.1)	
DLCST [51]	LDCT	Denmark, 1 centre	561	4104	2052	57.9 ± 4.8 (49–71)	1147/2052 (55.9)	1545/2052 (75.3)	507/2052 (24.7)	

	Control				2052	57.8 ± 4.8 (49–71)	1120/2052 (54.6)	1579/2052 (76.9)	473/2052 (23.1)	
Garg *et al.* 2002 [52]	LDCT	USA, 1 centre	304	239	92 (55 high risk, 37 medium risk)	68.1 ± 6.2 (high risk) 63.3 ± 6.6 (medium risk)	185/190	NR	NR	NR
	Control				98 (47 high risk, 51 medium risk)	67.4 ± 8.2 (high risk) 62.1 ± 7.6 (medium risk)		NR	NR	NR
ITALUNG [53]	LDCT	Italy, 3 centres (urban)	71,232 letters were sent. There were 17,055 (23.9%) responders	3206		Recruited: 55–69 years 55–59 years, *n* = 734 60–65 years, *n* = 580 > 65 years, *n* = 299	1035/1613 (32.28)	432/1406 (13.47)	146/1406 (4.55)	
	Control					Recruited: 55–69 years 55–59 years, *n* = 670	1039/1593 (32.41)	406/1593 (12.66)	148/1593 (4.62)	

						60–65 years, *n* = 626 > 65 years, *n* = 297				
Lung SEARCH [54]	LDCT	UK, 10 centres (urban)	NR	785	> 90% of screened subjects provided sputum in year 1	63 (mean age)	(52)	(56)	(44)	
	Control			783						
LUSI [55]	LDCT	Germany, (NR, but 5 study areas)	292,440	4052	2029	50–54 years, *n* = 942 55–59 years, *n* = 518 60–64 years, *n* = 344 65–69 years, *n* = 225	1315/2029 (64.8)	1259/2029 (62.1)	770/2029 (37.9)	NR
	Control				2023	50–54 years, *n* = 932 55–59 years, *n* = 528 60–64 years, *n* = 341	1307/2023 (64.6)	1247/2023 (64.6)	775/2023 (38.3)	NR

						65–69 years, *n* = 222				
MILD [57]	LDCT (annual)	Italy, 1 centre	4099	1190	1190	57	814/1190 (68.4)	820/1190 (68.9)	370/1190 (31.1)	
	LDCT (biannual)			1186	1186	58	813/1186 (68.5)	810/1186 (68.3)	376/1186 (31.7)	
	Control			1723	1723	57	1090/1723 (63.3)	1546/1723 (89.7)	177/1723 (10.3)	
NELSON [60]	LDCT	The Netherlands and Belgium, 4 centres	606,409	15,822	7915	58.0 (IQR 54.0–62.0)	6328/7582 (83.5)	4215/7582 (55.6)	3367/7582 (44.4)	
	Control				7907	57.0 (IQR 8.0)	6275/7453 (84.2)	4077/7434 (54.8)	3357/7434 (45.2)	377/7396 (4.7)
UKLS [63]	LDCT	UK, 2 centres	247,354 sent questionnaire; 8729 eligible	2028	1994	67 (67.1 ± 4.1)	1529/2028 (75.4)	777/2028 (38.3)	1249/2028 (61.6)	498/2028 (24.6)
	Control			2027	2027	67 (66.9 ± 4.1)	1507/2027 (74.3)	791/2027 (39.0)	1236/2027 (61.0)	554/2027 (27.3)
Yang [9]	LDCT	China, general practices Shanghai	Not reported	6657	3350	59.9 ± 5.8	1625(46.3)	777 (38)		
	Control						1489(47.3)	701(21.8)		

Table A2 Characteristics of study populations [LDCT vs. usual care (no screening)]

Study identifier	Arm	Country and number of centres	Number of patients approached	Number of patients randomised	Number of patients screened at baseline (*n/N*, %)	Characteristics of patients at baseline				
						Median age (years), (range)	Male, n (%) intervention/ control	Current smokers n (%) intervention/ control	Former smokers n (%) intervention/ control	Family history of lung cancer, n (%) intervention/ control
Depiscan [50]	LDCT	France, 14 centres	830	765	385	56 (47–75)	274/385 (71)	238/385 (65)	129/385 (35)	NR
	CXR				380	56 (47–76)	267/380 (70)	224/380 (64)	127/380 (36)	NR
LSS-PLCO [49]	LDCT	USA, 6 centres	653,417 mailed; 12,270 contacted; 4828 eligible	3318	1660	50–59 years, *n* = 616 60–64 years, *n* = 514 65–69 years, *n* = 337 70–74 years, *n* = 193	965/1660 (58.1)	961/1660 (57.9)	699/1660 (42.1)	NR
	CXR				1658	50–59 years, *n* = 624 60–64 years, *n* = 500 65–69 years, *n* = 3448	965/1658 (59.0)	947/1658 (57.1)	711/1658 (42.9)	NR

						70–74 years, *n* = 186				
NLST [62]	LDCT	USA, 33 centres	NR	53,454	26,722	< 55 years, *n* = 2 55–59 years, *n* = 11,440 60–64 years, *n* = 8170 65–69 years, *n* = 4756 70–74 years, *n* = 2353 > 74 years, *n* = 1	15,770/26,722 (59.0)	12,862/26,722 (48.1)	13,860/26,722 (51.8)	5815/26,723 (21.8)
	CXR				26,722	< 55 years, *n* = 4 55–59 years, *n* = 11,420 60–64 years, *n* = 8198 65–69 years, *n* = 4762 70–74 years, *n* = 2345 > 74 years, *n* = 3	15,762/26,732 (59.0)	12,900/26,722 (48.3)	13,832/26,732 (51.7)	5806/26,733 (21.7)[a]

						Missing, *n* = 1				

NR, not reported.

Table A3 Characteristics of study populations (LDCT vs. CXR)

Study identifier	Method of recruitment	Definition of high-risk individuals at baseline	Exclusion criteria	Initial adherence to screening
DANTE [43]	Via family doctors, large-scale mailings, media, internet, hospital boards and leaflets. Only males recruited	Aged 60–74 years; current or former smokers (≥ 20 pack-years; quit < 10 years before recruitment)	Other disease with < 5 years' life expectancy; < 5 years' disease-free laryngeal and non-melanoma skin cancer; treatment of other cancer in the last 10 years; unable to engage with follow-up protocol	Did not provide consent (post randomisation): 91/1403 vs. 166/1408
				Non-adherence to baseline screening:97% (1264/1300) vs. 96% (1186/1232)
				Proportion attending all five CT scans (of those with a baseline scan): 93% (1184/1264)
Depiscan [50]	performed across two study via family and occupational doctors (selection and enrolment);information was provided, consent obtained and randomisation performed across two study appointments. Males and females recruited	Aged 50–75 years; current or former smokers (≥ 15 cigarettes per day for ≥ 20 years; quit < 15 years before recruitment)	History of other cancer; disease that would hinder or prevent thoracic surgery or diagnostic procedure, including pulmonary infections; congestive heart failure/recent myocardial infarction; heavy exposure to asbestos; prior disease that may look radiologically similar to lung cancer; current symptoms	Non-adherence to baseline screening: 144 (19%) across both arms, significantly lower in the CT arm (55/385 vs. 89/380) and in older participants

Table A4 Methods of recruitment and adherence to screening regimens in trials comparing LDCT to no screening or CXR (continued)

DLCST [51]	Via local and regional media (free newspapers). Males and females recruited	Aged 50–70 years; current or former smokers (≥ 20 pack-years; quit at > 50 years of age and < 10 years before recruitment)	Other disease with < 10 years' life expectancy; history of treatment for lung or breast cancer, malignant melanoma, or hypernephroma; disease-free < 5 years for other cancers and <2 years for tuberculosis; CT scan ≤ 1 year ago; body weight > 130 kg; current symptoms; FEV2 of ≤ 30% of normal; not able to climb 36 steps without stopping	Non-adherence to baseline screening: low in both arms, higher in the CT arm (5/2052 vs. 0/2052). Mean participation rates across all study time-points: significantly higher in the CT arm (95.5% vs. 93.0%)
Garg *et al*. 2002 [52]	Via medical centre for veterans and associated clinics. Mostly males recruited	Aged 50–80 years; current or former smokers (≥ 30 pack-years)	Other disease with < 6 months' life expectancy; thoracic CT scan ≤ 3 years ago; pregnancy; not able to provide consent or engage with follow-up protocol	Adherence not reported
		High-risk group also had airflow obstruction diagnosed in a sputum cytology cohort study. Moderate-risk group were randomly selected but met above risk criteria	Moderate-risk group only: symptomatic COPD; airflow obstruction; non-compliance with inhalers	

Table A4 Methods of recruitment and adherence to screening regimens in trials comparing LDCT to no screening or CXR (continued)

ITALUNG [53]	Via letter from family doctors. Males and females recruited	Aged 55–69 years; current or former smokers (≥ 20 pack-years; quit < 10 years before recruitment)	History of other cancer (except non-melanoma skin cancer); unable to engage with follow-up protocol involving thoracic surgery	Adherence to baseline screening: 87% (1406/1613). Proportion attending four CT scans: 79%
LSS-PLCO [49]	Via large-scale mailings, clinician recommendations, media adverts and posters. Males and females recruited	Aged 50–74 years; current or former smokers (≥ 30 pack-years; quit < 10 years before recruitment	History of lung cancer; current treatment for other cancer (except non-melanoma skin cancer); thoracic or lung CT scan ≤ 2 years ago; previous lung resection; participation in other cancer trials (except smoking cessation)	Adherence to baseline screening: higher in CT arm, 96% (1586/1660) vs. 93% (1550/1658)
				Proportion attending at 1 year: higher in CT arm, 85.8% vs. 79.9%; adherence significantly lower in those with positive screens at baseline
LungSEARCH [54]	Via family doctors and hospital clinics. Males and females recruited	Current or former smokers (≥ 20 pack-years; smoked ≥ 20 years; quit < 8 years before recruitment); COPD	No history of cancer	Adherence to baseline screening not reported

Table A4 Methods of recruitment and adherence to screening regimens in trials comparing LDCT to no screening or CXR (continued)

LUSI [55]	Via large-scale mailings to participants identified through population registers in the local area. Males and females recruited	Aged 50–69 years; current or former smokers (≥ 15 cigarettes per day for ≥ 25 years or ≥ 10 cigarettes per day for ≥ 30 years; quit < 10 years before recruitment)	Other disease with < 10 years' life expectancy; cancer diagnosis ≤ 5 years ago, unable to engage with surgical treatment	Adherence to baseline screening: high (99.9%) in both arms 2028/2029 vs. 2022/2023, and similar in both arms across five screening rounds

MILD [57]	Via media (newspaper, television) adverts. Males and females recruited	Aged ≥ 49 years; current or former smokers (≥ 20 pack-years; quit < 10 years before recruitment)	History of cancer ≤ 5 years ago	Proportion attending at ≥1 CT scan: 97% in both screening groups (1149/1186 biennial; 1152/1190 annual)
				Proportion of participants adhering over
				Proportion of participants adhering over the study: 96.1% annual 95.1% biennial

Table A4 Methods of recruitment and adherence to screening regimens in trials comparing LDCT to no screening or CXR (continued)

NELSON [59,60]	Via population registries across two countries (the Netherlands and Belgium). Males only to start with, recruitment later expanded to females	Aged 50–75 years; current or former smokers (≥ 15 cigarettes per day for ≥ 25 years or ≥ 10 cigarettes per day for ≥ 30 years; quit < 10 years before recruitment)	Lung cancer diagnosis < 5 years ago or ≥ 5 years ago with current treatment; history of melanoma, hypernephroma, renal or breast cancer; history of other cancers (unless curatively treated > 5 years ago without recurrence); pneumonectomy; thoracic CT scan < 1 year ago; body weight ≥ 140 kg; moderate/bad health (self-report) and not able to climb two flights of stairs	Adherence to first screening round: 95.5% (7557/7915)
				Proportion attending at ≥ 1 CT scan: 95.8% (7582/7915)

NLST [6,62]	Via targeted mailings, media adverts (local radio and newspapers, television, websites, internet adverts), health fairs, unions, local branches of the American Cancer Society, and community groups. Recruitment included strategies to improve access to the study for minority groups. Males and females recruited	Aged 55–74 years; current or former smokers (≥ 30 pack-years; quit < 15 years before recruitment)	History of lung cancer; haemoptysis; thoracic CT scan < 18 months ago; unexplained weight loss (> 6.8 kg in last year)	Adherence to first screening round: high across both arms 98% (52,344/ 53,439), 98.5% (26,309/26,715) in the CT arm vs. 97.4% (26,035/26,724) in the control arm
Table A4 Methods of recruitment and adherence to screening regimens in trials comparing LDCT to no screening or CXR (continued)				
UKLS [63]	Via letter, sent by a data management company on behalf of the recipient's PCT. Letter recipients of the correct age, living in six PCTs around Liverpool and Cambridgeshire, were randomly selected using NHS PCT records. Males and females recruited	Aged 50–75 years; using the LLPv2 risk prediction model, ≥ 5% 5-year risk of lung cancer	Other disease that would prevent screening or lung cancer treatment; thoracic CT scan < 1 year ago; not able to lie flat; not able to provide consent	Adherence to baseline screening: 98.3% (1994/2028)

Yang China [9]	Through general practitioners and advertising leaflets in 6 communities (2 housing estates in each community were randomly selected, 6–8 residential buildings were randomly selected at each housing estate, ≥500 residents in these 6–8 buildings were invited to fill in a homebased questionnaire).	1) current or former smokers who had a history of at least 20 pack-years of cigarette smoking, and for former smokers, no more than 15 years since quitting; 2) cancer history of any kind in close family members; 3) cancer history of any kind for the participant; 4) occupational exposure to carcinogenic agents (asbestos, dust or radiation); 5) long history of passive smoking (>2 h every day in homes or indoor workplaces for at least ten years); and/or 6) long-term exposure to cooking oil fumes (cooking history of stir frying, frying or deep frying>50 dish-years).	Previously received a diagnosis of lung cancer, had a performance status (PS) >2, had a CT scan of chest within the last 12 months or had a diagnosis of any other cancer (including lung cancer) within the past 5 years.	LDCT was performed in 98.9% (3512/3550) of the participants in the LDCT group.
PCT, primary care trust.				

Table A4 Methods of recruitment and adherence to screening regimens in trials comparing LDCT to no screening or CXR

Study	CT technology (vendor CT scanner)	Multi or single detector	Voltage (kV)		Slice thickness (mm)	Volumetric analysis	Pitch	Estimated average effective dose (mSv)
DANTE [43]	NR	Multi after 2003 and single before 2003	140	40	5	NR	1.25	NR
DLCST [51]	Philips Mx 8000 (Philips Medical Systems, Eindhoven, the Netherlands)	Multi (16 slice)	120	40	1–1.5	Philips evaluation semiautomated software	1.5	1
Garg *et al.* [52]	NR	Single	120	50	NR	NR	02:01	NR
ITALUNG [53]	NR	1 × single and 4 × multi	120–140	20–43	1–1.25 for multislice	NR	1–2	NR
LungSEARCH [54]	NR	NR	NR	NR	NR	NR	NR	NR
LUSI [55]	Unspecified Toshiba and Siemens scanners, (switch of technology at 2010)	Multi (16 and 128 slice) after 2010 and single before 2010	NR	NR	1	Computer-aided detection (MEDIAN Technologies, Valbonne, France) with volumetric software	NR	1.6–2
MILD[57]	Somatom Sensation 16, Siemens (Siemens Medical Solutions, Forchheim, Germany)	Multi (16 slice)	120	30	1	LungCare, Siemens, semi-automated software (Siemens Healthcare,Forchheim, Germany)	1.5	NR

NELSON [59,60]	M×8000 IDT (Philips Medical Systems, Cleveland, OH, USA) or Brilliance 16P, Philips (Philips Medical Systems, Cleveland, OH,USA), or Sensation-16, Siemens (Siemens Medical Solutions,Forchheim, Germany)	Multi (16 slice)	80–90 (< 50 kg) 100 (< 60 kg) 120 (60–80 kg) 140 (> 80 kg)	20	1	LungCare, Siemens, semi-automated software	1.5	< 0.4 (< 60 kg) < 0.8 (60–80 kg) < 1.6 (> 80 kg)
UKLS [63]	Unspecified Siemens and Philips Brilliance 64 (Philips Medical Systems, Cleveland, OH, USA)	Multi (128 and 64 slice)	Automated based on BMI	Automated based on BMI	1	Siemens syngo LungCare, version Somaris/5 VB 10A, (Siemens Medical Solutions, Forchheim,Germany)	0.9–1.1	NR

Table A5 Computed tomography parameters for LDCT vs. CXR (continued)

Study	CT technology (vendor CT scanner)	Multi or single detector	Voltage (kV)	Tube current-time product (mAs)	Slice thickness (mm)	Volumetric analysis	Pitch	Estimated average effective dose (mSv)
Depiscan [50]	NR	Multi	100–140 automated based on BMI	20–100 automated based on BMI	1.25–3	NR	NR	NR
LSS-PLCO [49]	Variable and not specified	Multi (inclusion criteria said must have a history of a spiral/helical CT scan)	120–140	60	NR	NR	2	NR
NLST [6,62]	97 different scanners	Multi > 4 slices	120–140	40–80 automated based on BMI	1–2.5	NR	1.25–2 (typically 1.5)	1.5
Yang China [9]	64-detector row scanner (Brillance, Philips, USA)	64 slices	140,	40	5	NR	1.25	NR
BMI, body mass index; LSS-PLCO, Lung Screening Study as part of the Prostate, Lung, Colorectal and Ovarian cancer screening trial; NR, not reported								

Table A5 Computed tomography parameters for LDCT vs. CXR

Study	Random sequence generation (selection bias) and support for judgement	Allocation concealment (selection bias) and support for judgement	Blinding of participants and personnel (performance bias) and support for judgement	Blinding of outcome assessment (detection bias) and support for judgement	Incomplete outcome data (attrition bias) and support for judgement	Selective reporting (reporting bias) and support for judgement
Lung cancer and overall mortality						
DANTE [43]	Low 'Subjects were randomised by a 1 : 1 scheme in blocks of four and stratified by centre according to a computer-generated list supplied by the data centre each week before the enrolment sessions'	Unclear No statement found	Low Due to trial design no blinding possible but bias for these endpoints is unlikely.	Unclear 'Copies of any medical records concerning the underlying diagnosis or death cause are requested…' and 'A panel blinded to patient's assignment reviewed the clinical cases whenever several competing causes of death were possible'	Low 'Life status data and death certificates were requested for the entire study population from local health registries' and , the cause of death was known through death certificates alone or remained unknown in two subjects.'	Low Outcome prespecified
DLCST [51]	Low 'Participants were randomized by a computer program	Unclear No statement found	Low Due to trial design no blinding possible but	Low 'An international independent death	Low Vital status was unknown in 20/2052 in the	Low 'The primary outcome was assessment of lung cancer mortality and all-cause mortality in the two groups.'

	(random permuted blocks of 10 participants) to either annual screening by low-dose computed tomography (the screening group) or the control group, which was not offered CT screening.'		bias for these endpoints is unlikely.	review board will be established.'	LDCT and 14/2052 in controls due to emigration.	
LUSI [55]	Low 'Electronic randomize-ation was carried out using the randomizat-ion tool RANDI developed in the Biostatistic Branch of the DKFZ. The block randomization was stratified by age …, gender, and smoking … status.	Unclear No statement found	Low Due to trial design no blinding possible but bias for these endpoints is unlikely.	Low 'An end point committee composed of a chest surgeon, two radiologists and a pathologist classified the cases using methods identical to those in NELSON, with full blinding with regard to the allocation of patients.' to either the screening or control arm.	Low 'On April 30, 2018, after an average observation time of 8.8 years, 3,741 subjects were documented to be still alive, whereas and 298 had deceased; 13 were lost to follow-up.'	Low Primary outcome not stated but bias for mortality data is low.
Mild [57]	low	Moderate	Low	Low	Low	Low

	Centralized stratified randomization was accomplished by the use of blocks of variable size. The list of randomization was stratified by reference center, age …, and duration of smoking . The group randomized to receive LDCTwas further randomized to receive LDCTevery 12 months (annual) or every 24 months (biennial).	No statement found Imbalance for important baseline characteristics between treatment arms	Due to trial design no blinding possible but bias for these endpoints is unlikely.	Outcome assessment blinded-. Cancer Registry Office database of Lombardy which traced the vital status of all participants blindly, without knowing the random allocation' The cause of death was missing in 4.1% of participants (10/243, 3 in intervention and 7 in control arm).	For deceased participants, we obtained the death certificate from the Istituto Nazionale di Statistica (ISTAT). 'Among the 3856 survivors, 93.5% (n=3607) of participants reached the 9 years of follow-up and 71% (n=2739) accumulated 10 years of follow-up. Only one subject was lost to follow-up.'	Outcomes prespecified
Nelson [59,60]	Unclear No details randomisation methods found	Low 'We randomly assigned eligible participants in NELSON, who were recruited as described previously, 11 to undergo CT screening at baseline (first	Low Due to trial design no blinding possible but bias for these endpoints is unlikely.	Low 'All relevant medical information will be collected and blinded for the participant's identity and study arm by an individual	Low 'Follow-up data were retrieved from national linkages at approximately 5, 7, and 10 to 11 years of complete	Low Outcome prespecified overall mortality not reported in females.

	No information found	round), 1 year later (second round), and 3 years later (third round, 2 years after the second round), or no screening.'		who is not otherwise involved in the trial.' And 'A clinical expert committee was formed to assign the cause of death by an evaluation process using a flow chart and predeter-mined criteria.'	follow-up. A total of 18 persons (13 men and 5 women) could not be linked, because a digital consent form could not be retrieved.'	
NLST [62]	Low 'Randomization occurred after data co-ordinating centres confirmed that eligibility criteria had been met for a given individual; participants were then assigned to either the computerized tomography arm or chest radiograph arm in a 1 : 1 ratio, stratifying by site, sex, and 5-year age group. Stratified randomization was	Unclear No statement found	Low Due to trial design no blinding possible but bias for these endpoints is unlikely.	Low Outcome assessment blinded 'An endpoint verification team determined whether the cause of death was lung cancer . . . members of the team were not aware of the group assignments'	Low Primary outcome assessed via with of use of National Death Index. Cancer status unknown of 67/26309 in LDCT and 60/26035 in CXR groups	Low Outcomes prespecified

	accomplished by use of a block size of six or eight, with block size chosen at random'.					
UKLS [70]	Low 'Recruits were randomised by computer into the intervention arm (LDCT scan, screen group) or the control arm (usual care, non-screen group) at a ratio of 1 : 1'.	Low 'Each subject was given a randomly generated unique code, consisting of eight characters (0–9; A–Z). All subjects who had given fully informed signed consent were available for randomisation; a minimum of two subjects from the same site (Liverpool or Papworth) were required in order for randomisation to be implemented. Each time randomisation took place, the computer generated a random shift number so that the order of characters in each participant's unique code was shifted by this number. The rearranged codes were then ordered alphanumerically, and split on a 1 : 1 basis into A	Low Due to trial design no blinding possible but bias for these endpoints is unlikely	Low Outcomes from UK cancer and death registry data were provided by NHS Digital and the National Cancer Registration and Analysis Service (NCRAS) who were not aware of the participants' allocated trial arm.	Low Mortality data for lung cancer and overall mortality could not be linked for 41 and 46 individuals in the LDCT and control groups	Low Outcomes prespecified

		(intervention) or B (control) groups. ‘				
Psychological consequences of screening and HRQoL						
DLCST [51]	Low ‘Participants were randomised by use of an in-house computer program developed by Asger Dirksen, M.D.,D.Msc. (random permuted blocks of 10 participants) to either annual screening by low-dose computed tomography (the screening group) or the control group, which was not offered CT screening’	Unclear No statement found	High Due to trial design no blinding possible.	High Outcome assessment was self-reported and reporters were unblinded	High 1845/2051 (89.9%) in LDCT and 1374/2052 (67.0%) answered at last follow-up questionnaire	Low Outcomes prespecified
NELSON [60]	Unclear No details on randomisation methods found	Unclear Sampling was done to obtain subset of all trial entrants taking part in HRQoL part of trial. The control arm was further sampled to obtain follow-up	High Due to trial design no blinding possible.	High Outcome assessment was self-reported and reporters were unblended.	High ‘The questionnaire response at T0 was 89.8% in the screen group and 85.9% in the control	Low Outcomes prespecified

		questionnaire data. Sampling said to be random			group; at T1 it was 87.7% (screen group only) and at T2 it was 89.3% in the screen group and 64.7% in the control group.'	
UKLS [63]	Low 'Recruits were randomised by computer into the intervention arm (LDCT scan, screen group) or the control arm (usual care, non-screen group) at a ratio of 1 : 1'.	Low 'Each subject was given a randomly generated unique code, consisting of eight characters (0–9; A–Z). All subjects who had given fully informed signed consent were available for randomisation; a minimum of two subjects from the same site (Liverpool or Papworth) were required in order for randomisation to be implemented. Each time randomisation took place, the computer generated a random shift number so that the order of characters in each participant's unique	High Due to trial design no blinding possible.	High Outcome assessment was self-reported and reporters were unblended.	High 1553/2028 (82.3%) in LDCT and 1302/2027 (65.3%) completed questionnaires at T2.	Low Outcomes prespecified

		code was shifted by this number. The rearranged codes were then ordered alphanumerically, and split on a 1 : 1 basis into A (intervention) or B (control) groups. ‘				
Smoking cessation						
DLCST [81]	Low ‘Participants were randomised 1 : 1 to annual LDCT (screening group) or no CT screening (control group). Permuted block randomisation with fixed block size (10) was used.’	Unclear No statement found	High Due to trial design no blinding possible.	High ‘Smoking habits (self-reported) were assessed annually, starting with an initial (baseline) screening and followed by four annual screening rounds, giving a total of five screening years. At baseline and the second screening visit, self-reported smoking habits were objectively verified by measuring carbon	High ‘Missing recording of smoking status was mainly due to loss to follow-up increased from 8.6% at baseline to 14.6% at the final screening round (year 5), and were more frequent in the control group (12%) than the CT group (6%; $p<0.001$, $\chi 2$ test).’	Unclear Not prespecified

				monoxide (CO) levels in exhaled breath.'		
UKLS [82]	Low ,High-risk individuals who gave informed written consent were randomised on a 1:1 ratio to the intervention (screening) or control arms.'	Low 'Randomisation used unique random personal ID codes and computer-generated sequencing for allocation concealment.'	High Due to trial design no blinding possible.	High 'Smoking cessation was assessed using self-report at T1 and T2.'	High 'T2 completion rates were n=488/749 (65%) for the screening arm and n=377/775 (49%) for the control arm (total T2 n=865).'	Unclear Psychosocial variables and smoking cessation were a focus of the trial.

Table A6 Risk of Bias assessment with support for judgement for critical and important outcomes that were available in intervention and control arms for trials comparing LDCT with no screening or CXR

			Mortality								
	Comparator		Lung cancer					All-cause			
Study identifier	Usual care		Number of events in the LDCT group	Total number of participants in the LDCT group	Number of events in the control group	Total number of participants in the control group		Number of events in the LDCT group	Total number of participants in the LDCT group	Number of events in the control group	Total number of participants in the control group
DANTE [71]	Usual care		59	1264	55	1186		180	1264	176	1186
DLCST [217]	Usual care		39	2052	38	2052		165	2052	163	2052
MILD (annual & biannual) [57]	Usual care		40	2376	40	1723		137	2376	107	1723
NELSON	Usual care		181*	7900*	242*	7892*		868±	6583±	860±	6612±
NLST [6]	CXR		356	26722	443	26732		1877	26722	2000	26732
UKLS [70]	Usual care		30	2028	46	2027		246	2028	266	2027
*Males and females, ±males only reported											

Table A7 Mortality data for lung cancer and all causes in trials with ≥5 years of follow-up

Study identifier	Method of recruitment	Definition of high-risk individuals at baseline	Exclusion criteria	Initial adherence to screening
Czech [98]	Via general health examination of middle-aged males only	Aged 40–64 years; current smokers (approximate lifetime consumption > 150,000 cigarettes)	History of pulmonary disease. Likely inability to participate over 3 years due to serious disease or other reasons	Adherence to screening over 3 years in screening arm 92.5% vs. control arm 94.7%
Mayo [99]	Via 'smoking survey' completed by outpatients at a general medical examinations by the Mayo Clinic	Aged > 45 years; current or former smokers (at least one pack per day at time of recruitment or within previous year)	History of known or suspected cancer of the respiratory tract (except roentgenographically occult cancer); < 5 years' life expectancy; unable to tolerate pulmonary resection; failure to complete general medical examination; insufficient mental capacity for study cooperation	Adherence to testing schedule over 6 years of screening averaged 75%
	If questionnaire categorised as 'high-risk' males only were referred to the study			
PLCO (for sensitivity analysis only) [100]	Via mass mailing of general population. A subset of entire PLCO population in line with population characteristics of NLST were used for this analysis. Males and females recruited	Aged 55–74 years; current or former smokers (≥ 30 pack-years; quit < 15 years before recruitment)	History of prostate, lung, colorectal or ovarian cancer, or current cancer treatment or removal of one lung	Adherence to baseline screening, screening arm 85.9% (13,035/15,183)
				Overall adherence to expected screens, screening arm 81.4% (48,330/15,183)

Table A8 Characteristics of recruitment and adherence in trials comparing CXRs to no screening

Study identifier (country)	Screening programme comparison	Definition of a positive scan for lung cancer	Imaging evaluation and interpretation strategy	Diagnostic follow-up for suspicious abnormality finding
Kubík et al. (Czech Republic) [98]	CXR	CXR	CXR	CXR
	(at baseline, 6-monthly during years 1–3, and then at years 3, 4 and 5 and 6, screening also included sputum cytology testing)	Positive if abnormality identified (reader decision whether or not further investigation was required)	Chest photofluorogram, posteroanterior view	Follow-up protocol
		Other	Double-reading by chest physician and chest radiologist. Decision based on consensus (third experienced physician arbitrated disagreements)	Positive CXR – referral to specialist diagnostic hospital ward (if sputum signs – recommendation for inpatient stay), fibre-optic animation, additional CXR, (including whole-lung CXR), otorhinolaryngological examination (for exclusion purposes)
		Also sent for further investigation if one or more of the following was evident: patient approached with symptoms, cancer or atypical cells from sputum testing, bloody sputum		
	No screening			
	(+ CXR at baseline, years 3, 4, 5 and 6, included sputum cytology testing at same times as CXR)			
Mayo (USA) [99]	CXR	CXR	CXR	CXR
	(4-monthly, screening also included sputum cytology testing, medical history review)	Not clear	Stereo chest roentgenograms, standard size	Follow-up protocol
			Double-reading by chest physician and radiologist. Decision based on consensus (another chest physician arbitrated disagreements)	Positive CXR, suggesting lung cancer – review of clinical data

				Positive CXR, new or growing abnormality – work-up could include additional CXR and sputum testing, bronchoscopy (with or without fluoroscopic guidance)
	Usual care			
	(annual CXR and sputum cytology testing)			
PLCO (USA [100]	CXR	CXR	CXR	CXR
	(at baseline, annually up to 4 years)	Positive if the readers felt that one of the following was evident and suspicious: any nodule, mass, infiltrate or other abnormality	Posteroanterior CXR	No study follow-up protocol, positive CXR follow-up was decided by patients and their health-care providers
	No screening			

Table A9 Characteristics of screening programmes: CXRs

						Characteristics of patients at baseline				
Study identifier	**Arm**	**Country; number of centres**	**Number of patients approached**	**Number of patients randomised**	**Number of patients screened at baseline**	**Age (years)**	**Male, n/N (%)**	**Current smokers, n/N (%)**	**Former smokers, n/N (%)**	**Family history of LC, n/N (%)**
Czech[98]	CXR	Czech Republic; six districts	6364	6346	3172	40–44: n = 487	3172/3172 (100)	3172/3172 (100)	NR	NR
						45–49: n = 716				
						50–54: n = 923				
						55–59: n = 582				
						60–64: n = 464				
	Control				3174	40–44: n = 499	3174/3174 (100)	3174/3174 (100)	NR	NR
						45–49: n = 710				
						50–54: n = 926				
						55–59: n = 584				
						60–64: n = 455				
Mayo[99]	CXR	USA; NR	NR	10,933 screened; 9211 randomised	4618	< 50: n = 1159	4618/4618 (100)	NR/NR (90)	NR/NR (10)	NR
						50 to < 55: n = 1102				
						55 to < 60: n = 1042				
						60 to < 65: n = 811				
						65 to < 70: n = 483				
						≥ 70: n = 21				
	Control				4593	< 50: n = 1154	4593/4593 (100)			NR
						50 to < 55: n = 1135				
						55 to < 60: n = 1019				

						60 to < 65: n = 784				
						65 to < 70: n = 469				
						≥ 70: n = 32				
PLCO (for sensitivity analysis only)[100]	CXR	USA; 10 centres	154,901	77,445	15,183	55 to 59: 25 850 (33.4) 60 to 64: 23 784 (30.7 65 to 69: 17 457 (22.5) 70 to 74: 10 354 (13.4)	9252/15,183 (60.9)	6146 (40.5)	32 555 (42.0)	7930 (10.6)
	Control			77,456	15,138	55 to 59:)25 839 (33.4) 60 to 64: 23 773 (30.7) 65 to 69: 17 473 (22.6) 70 to 74: 10 371 (13.4)	17 473 (22.69110/15,138 (60.2)	6069/15,138 (40.1)a	32 136 (41.5)	7729 (10.5)
LC, lung cancer; NR, not reported. a Calculated from raw data, differs from PLCO107 for which % is reported as 40.3%.										

Table A10 Characteristics of study populations (CXR vs. usual care)

Study	Random sequence generation (selection bias) and support for judgement	Allocation concealment (selection bias) and support for judgement	Blinding of participants and personnel (performance bias) and support for judgement	Blinding of outcome assessment (detection bias) and support for judgement	Incomplete outcome data (attrition bias) and support for judgement	Selective reporting (reporting bias) and support for judgement
Lung cancer mortality						
Czech[98 218]	Unclear The computer program to allocate the study participants to either of two study groups was done so that the groups did not significantly differ as to age, lifetime number of cigarettes smoked thus far, socioeconomic group, occupational exposure to noxious pollutants, and place of residence (district)	High Subjects with no evidence of lung cancer on initial screening were randomized to either of two study groups: (1) experimental or close-surveillance group, in which tests were to be repeated every 6 months during a total period of three years; and (2) control group that had the final screening 3 years after the initial examination. No baseline characteristics provided by study arm.	Unclear Due to trial design no blinding possible but bias for these endpoints is unlikely.	Unclear Not mentionend	High Numbers of individuals with loss to follow-up or unknown cancer status not mentioned	Low Outcome prespecified

Mayo[99 219 220]	Unclear No information provided	Low Those whose initial screens were considered satisfactory and negative for lung cancer (and who also met certain other qualifications) were randomly assigned either to a study group, which was offered rescreening every 4 months for 6 years, or to a control group, which was not offered any systematic rescreening but was advised to receive chest radiographs and sputum cytology testing at least once a year..'	High Contamination of the control population by nonstudy chest roentgenography was substantial. Fifty-five percent of the control group received chest radiographs during the final year of the MLP and 73% received them during the final 2 years.	Unclear The National Death Index (NDI; Hyattsville, MD) was used to follow-up the 6523 MLP participants who were known to be alive on July 1, 1983, and for whom vital status and date and cause of death, as of December 31, 1996, were unknown. No information on blinding provided.	High The vital status of 2669 MLP participants … was available from the Mayo Clinic's records. Information on 6523 of the remaining …was sent to the NDI. Of the 6523 records, no match was obtained for 1590, and no true match was obtained for 1972. These men were, therefore, assumed to be alive …'.	Low The primary outcome was assessment of lung cancer mortality'
PLCO (for sensitivity analysis only)[100]	Low Individuals who meet the eligibility criteria are randomized	low Individual randomization to either the intervention or usual care group was within	Low Due to trial design no blinding possible but	Low 'All deaths with causes potentially related to a PLCO	Unclear No information found on lung cancer mortality	Low The primary outcome was assessment of lung cancer mortality'.

	individually into intervention and control arms. The randomization scheme uses blocks of random permutations of varying lengths and is stratified by SC, gender, and age. Random assignment is implemented using compiled software and encrypted files loaded on SC microcomputers.	blocks that were stratified by screening center, sex, and age.	bias for these endpoints is unlikely.	cancer were reviewed, including any for which the participant had a prostate, lung, colorectal, or ovarian cancer or possible metastasis from 1 of these cancers and any of unknown or uncertain cause. Death reviewers were blinded to the trial group of the deceased participant.'	from flow chart in those who refused screening or left study.	

Table A11 Risk of bias with support for judgement for critical outcomes that were available in intervention and control arms in trials comparing CXR with no screening

24.Appendix 3 – Economic literature search strategies

Medline via Ovid (Ovid MEDLINE(R) ALL), date of search: 17.12.2020; date parameters: 1946 to December 15, 2020; 861 hits

---

1. exp Lung Neoplasms/

2. ((lung$ or bronch$ or pulmon$) adj3 (cancer$ or neopla$ or tumor$ or tumour$ or carcinoma$ or adenocarcinoma$ or small cell or squamous)).ti,ab,ot,kw.

3. (NSLC or NSCLC or SLC or SCLC).ti,ab,ot,kw.

4. 1 or 2 or 3

5. exp Tomography, X-Ray Computed/ or exp Radiography, Thoracic/

6. ((CT or CAT) adj3 (scan$ or screen$)).ti,ab,ot,kw. or (x ray or xray or x-ray or CXR or radiograph$).ti,ab,ot,kw.

7. ((computer$ adj3 tomogra$) and (scan$ or screen$)).ti,ab,ot,kw.

8. (tomogra$ or helix or helical or spiral$ or spiro$).ti,ab,ot,kw.

9. 5 or 6 or 7 or 8

10. 4 and 9

11. exp Economics/

12. Economics, Medical/

13. Economics, Nursing/

14. Economics, Pharmaceutical/

15. exp Economics, Hospital/

16. (economic$ or cost or costs or costly or costing or price or prices or pricing or priced or discount or discounts or discounted or discounting or ration$ or expenditure or expenditures or expense or expenses or financial or finance or finances or financed or budget$ or afford$ or pharmacoeconomic$ or pharmaco-economic$).ti,kf. or (economic$ or cost or costs or costly or costing or price or prices or pricing or priced or discount or discounts or discounted or discounting or ration$ or expenditure or expenditures or expense or expenses or financial or finance or finances or financed or budget$ or afford$ or pharmacoeconomic$ or pharmaco-economic$).ab. /freq=2

17. exp "Fees and Charges"/

18. (fee or fees or charge$ or preference$).tw.

19. (fiscal or funding or financial or finance).tw.

20. exp "Costs and Cost Analysis"/

21. exp Health Care Costs/

22. (cost* adj2 (effective* or utilit* or benefit* or minimi* or analy* or outcome or outcomes)).ab,kf.

23. (value adj2 (money or monetary)).ti,ab,kf.

24. exp Decision Support Techniques/

25. exp Models, Economic/

26. economic model*.ab,kf.

27. markov$.ti,ab,ot,kf.

28. Markov Chains/

29. monte carlo.ti,ab,ot,kf.

30. Monte Carlo Method/

31. (decision$ adj2 (tree$ or analy$ or model$)).ti,ab,ot,kf.

32. exp Decision Theory/

33. (survival adj3 analy$).tw.

34. "Deductibles and Coinsurance"/

35. exp Health Expenditures/

36. Uncertainty/

37. exp Budgets/

38. or/11-37

39. exp technology assessment, biomedical/ or (technology assessment* or HTA or HTAs or technology overview* or technology appraisal*).ti,ab,kf,kw.

40. 38 or 39

41. exp Animals/ not humans.sh.

42. 40 not 41

43. (2016* or 2017* or 2018* or 2019* or 2020* or 2021*).dt,ez,ed.

44. 10 and 42 and 43

Note: We added the HTA filter part from the CADTH filter "Systematic Reviews/Meta-Analysis/Health Technology Assessment – OVID Medline, Embase, PsycINFO" from https://www.cadth.ca/resources/finding-evidence/strings-attached-cadths-database-search-filters

Embase via Ovid, date of search: 17.12.2020; date parameters: 1974 to 2020 December 15; 2514 hits

1. exp lung cancer/ or exp lung tumor/ or exp bronchus cancer/

2. ((lung$ or bronch$ or pulmon$) adj3 (cancer$ or neopla$ or tumor$ or tumour$ or carcinoma$ or adenocarcinoma$ or small cell or squamous)).ti,ab,ot,kw.

3. (NSLC or NSCLC or SLC or SCLC).ti,ab,ot,kw.

4. 1 or 2 or 3

5. exp computer assisted tomography/

6. ((CT or CAT) adj3 (scan$ or screen$)).ti,ab,ot,kw.

7. ((computer$ adj3 tomogra$) and (scan$ or screen$)).ti,ab,ot,kw.

8. (tomogra$ or helix or helical or spiral$ or spiro$).ti,ab,ot,kw. or (x ray or xray or x-ray or CXR or radiograph$).ti,ab,ot,kw.

9. 5 or 6 or 7 or 8

10. 4 and 9

11. Economics/

12. Cost/

13. exp Health Economics/

14. Budget/

15. budget*.ti,ab,kw.

16. (economic* or cost or costs or costly or costing or price or prices or pricing or priced or discounted or discounting or ration* or pharmacoeconomic* or pharmaco-economic* or expenditure or expenditures or expense or expenses or financial or finance or finances or financed or budget* or afford*).ti,kw.

17. (economic* or cost or costs or costly or costing or price or prices or pricing or priced or discounted or discounting or ration* or pharmacoeconomic* or pharmaco-economic* or expenditure or expenditures or expense or expenses or financial or finance or finances or financed or budget* or afford*).ab. /freq=2

18. (fee or fees or charge$ or preference$).tw. or (fiscal or funding or financial or finance).tw. or (cost* adj2 (effective* or utilit* or benefit* or minimi* or analy* or outcome or outcomes)).ab,kw.

19. (value adj2 (money or monetary)).ti,ab,kw.

20. Statistical Model/

21. economic model*.ti,ab,kw.

22. Probability/

23. markov*.ti,ab,kw. or exp Markov chain/

24. monte carlo.ti,ab,kw. or monte carlo method/

25. (survival adj3 analy*).tw.

26. Decision Theory/

27. Decision Tree/

28. (decision* adj2 (tree* or analy* or model*)).ti,ab,kw.

29. or/11-28

30. exp biomedical technology assessment/ or (technology assessment* or HTA or HTAs or technology overview* or technology appraisal*).ti,ab,kw.

31. 29 or 30

32. 10 and 31

33. (2016* or 2017* or 2018* or 2019* or 2020* or 2021*).dp,dc,dd.

34. 32 and 33

Note: We added the HTA filter part from the CADTH filter "Systematic Reviews/Meta-Analysis/Health Technology Assessment – OVID Medline, Embase, PsycINFO" from https://www.cadth.ca/resources/finding-evidence/strings-attached-cadths-database-search-filters

Web of Science Core Collection via Clarivate Analytics, date of search: 17.12.2020; date parameters: 2016-2020; 1436 hits

Database: Web of Science Core Collection: Science Citation Index Expanded (SCI-EXPANDED) and Conference Proceedings Citation Index –Science (CPCI-S).

1. TS=((lung* or bronch* or pulmon*) near/3 (cancer* or neopla* or tumor* or tumour* or carcinoma* or adenocarcinoma* or “small cell” or squamous))
2. TS=(NSLC or NSCLC or SLC or SCLC)
3. TS=((CT or CAT) near/3 (scan* or screen*))
4. TS=((computer* near/3 tomogram*) and (scan* or screen*))
5. TS=(tomogra* or helix or helical or spiral* or spiro* or x ray or xray or x-ray or CXR or radiograph*)
6. #1 or #2
7. #3 or #4 or #5
8. #6 and #7
9. TS=((pharmacoeconomic* or socioeconomics or economic* or pric* or cost* or cba or cea or cua or “health utilit*” or “value for money”))
10. TS=(economic* or price or prices or pricing or priced or discount or discounts or discounted or discounting or ration* or expenditure or expenditures or budget* or afford* or pharmacoeconomic* or pharmacoeconomic* or fee or fees or charge* or preference* or fiscal or funding or financial or finance or economic model* or markov* or monte carlo)
11. TS=(cost* near/2 (effective* or utilit* or benefit* or minimi* or analy* or outcome or outcomes))
12. TS=(value near/2 (money or monetary))
13. TS=(decision near/2 (tree* or analy* or model*))
14. TS=(survival near/3 analy*)
15. #9 or #10 or #11 or #12 or #13 or #14
16. #15 and #8

Note: TS searches in the Abstract, Title, and/or Keywords fields of a record; Indexes=SCI-EXPANDED, CPCI-S Timespan=2016-2020.

25.Appendix 4 – Quality assessment using the CHEC-checklist

Author	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19
Chirikos 2003 #	N	Y	N	N	Y	Y	Y	Y	Y	N	Y	N	Y	Y	N	Y	N	N	N
Pyenson 2012 #	Y	Y	Y	N	Y	Y	Y	Y	Y	N	Y	Y	Y	N	N	U	N	N	N
Villanti 2013 #	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	N	N	N	N	N	N	N
Pyenson 2014 #	Y	Y	Y	N	Y	Y	Y	Y	Y	N	Y	Y	Y	N	N	Y	N	N	N
Tabata 2014 #	Y	Y	Y	N	U	U	U	U	U	N	Y	Y	Y	U	N	Y	N	N	N
Cressmann 2017	U	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	U	Y
Marshall 2001 #	N	Y	N	N	N	N	Y	Y	Y	Y	Y	U	Y	Y	N	Y	Y	N	N
Wisnivesky 2003 #	Y	Y	Y	N	U	Y	Y	Y	N	N	Y	N	Y	Y	N	Y	Y	N	N
Whynes 2008 #	Y	Y	Y	N	U	Y	Y	Y	Y	Y	Y	N	Y	Y	N	Y	Y	U	Y
Goulart 2012 #	Y	N	N	N	N	Y	Y	Y	N	N	Y	N	Y	N	N	Y	N	Y	N
Shmueli 2013 #	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	U	N
Black 2014 #	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N	N	N
Black 2015 #	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N	N	N
HTA Field 2016 #	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	N	Y	Y	U	N
HTA Snowsill 2018	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	U	Y
Veronesi 2020	Y	Y	Y	U	N	Y	Y	Y	U	Y	Y	U	Y	U	Y	Y	Y	Y	U
Diaz 2021	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	U	Y	U
Wade 2018	U	Y	Y	U	U	Y	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	Y	Y	Y
Mahadevia 2003 #	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	U	N

Manser 2004 #	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	Y	Y	N	U	N
Yang 2017	Y	N	Y	Y	U	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Hofer 2018	Y	Y	Y	Y	U	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	U	Y	Y
McLeod 2018	Y	Y	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	U	U	Y	Y	Y	Y	Y
Sun 2021	Y	Y	Y	Y	N	Y	U	Y	Y	Y	Y	Y	Y	U	N	Y	U	Y	Y
Yuan 2021	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	U	Y	U	N	N
Esmaeili 2021	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	U	Y	Y	U	Y	Y	Y	Y
Gómez-Carballo 2021	Y	Y	Y	Y	Y	U	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	U
Kim 2021	Y	N	Y	U	Y	U	U	N	U	Y	Y	Y	Y	Y	Y	Y	U	Y	Y
Goffin 2016 #	Y	Y	Y	Y	Y	U	U	U	U	Y	Y	U	N	Y	N	Y	N	U	N
Jaine 2018	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	Y	Y	Y	U	Y
McMahon 2011 #	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N	U	N
Goffin 2015 #	Y	Y	Y	Y	Y	Y	U	U	U	Y	Y	U	N	Y	N	Y	Y	N	N
HTA Ontario 2014	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	U	N
Evans 2016	Y	Y	U	Y	U	Y	U	Y	Y	Y	U	U	Y	Y	Y	Y	Y	Y	Y
Ten Haaf 2017 #	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	Y	N	U	N
Treskova 2017	Y	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	Y	Y	Y	U	Y	Y	Y	U
Tomonaga 2018	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	U
Toumazis 2019	Y	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	Y	Y	Y	Y	Y	U	Y	U
Du 2020	Y	Y	Y	Y	U	Y	Y	Y	Y	Y	Y	U	Y	Y	Y	Y	Y	Y	Y
Griffin 2020	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Criss 2019	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	Y	U	Y	U	U	Y

Kumar 2018	Y	Y	Y	Y	U	Y	Y	Y	Y	Y	Y	Y	Y	U	U	Y	U	Y	Y
Hinde 2018	Y	U	Y	Y	U	Y	Y	Y	Y	Y	Y	U	Y	Y	U	Y	Y	Y	Y

[#] Study included in the HTA published by Snowsill et al. in 2018.
Consensus on Health Economic Criteria (CHEC)-list for economic evaluations
Q1 Is the study population clearly described?
Q2 Are competing alternatives clearly described?
Q3 Is a well-defined research question posed in answerable form?
Q4 Is the economic study design appropriate to the stated objective?
Q5 Is the chosen time horizon appropriate to include relevant costs and consequences?
Q6 Is the actual perspective chosen appropriate?
Q7 Are all important and relevant costs for each alternative identified?
Q8 Are all resources measured appropriately in physical units?
Q9 Are resources valued appropriately?
Q10 Are all important and relevant outcomes for each alternative identified?
Q11 Are all outcomes measured appropriately in physical units?
Q12 Are outcomes valued appropriately?
Q13 Is an incremental analysis of costs and outcomes performed?
Q14 Are all future costs and outcomes discounted appropriately?
Q15 Are all important variables, whose values are uncertain, appropriately subjected to sensitivity analysis?
Q16 Do the conclusions follow from the data reported?
Q17 Does the study discuss the generalisability of the results to other settings and patient/client groups?
Q18 Does the article indicate that there is not potential conflict of interest of study researcher(s) and funder(s)?
Q19 Are ethical and distributional issues discussed appropriately?

Judgements
N, No; U, Unclear; Y, Yes.

Table A12 Quality assessment using the CHEC-checklist

26.