Low-dose CT screening for lung cancer

Scope

Publication date
October 2020

Version
4.0

Authors (alphabetical order)
Prof. Dr. med. Heiner C. Bucher, MPH
Dr. David Shaw, MA, MSc, MML
Dr. Yuki Tomonaga, MSc

Head of institute
Prof. Dr. med. Heiner C. Bucher, MPH
Acknowledgment

The assessment team thanks the clinical and health economic experts for their support.
# Table of contents

Acknowledgment .......................................................................................................................... 2
Table of contents .......................................................................................................................... 3
Abbreviations ............................................................................................................................... 4

1 Background ............................................................................................................................... 5
2 Aims of the Health Technology Assessment report ................................................................. 5
3 Clinical effectiveness .................................................................................................................. 6
   3.1 Existing literature .................................................................................................................. 6
   3.2 Aim ..................................................................................................................................... 8
   3.3 Methods .............................................................................................................................. 8
4 Health economic evaluation ......................................................................................................... 15
   4.1 Aim ..................................................................................................................................... 15
   4.2 Pre-review of the health economic literature ...................................................................... 15
   4.3 Brief overview of the identified cost-effectiveness analyses ............................................. 16
   4.4 Approach to the health economic assessment ................................................................. 23
5 Ethical considerations .................................................................................................................. 27
   5.1 Context ............................................................................................................................... 27
   5.2 Ethical, legal and social issues ............................................................................................ 27
   5.3 Methods .............................................................................................................................. 27
6 References .................................................................................................................................. 29
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALCA</td>
<td>Anti-Lung Cancer Association</td>
</tr>
<tr>
<td>AUD</td>
<td>Australien Dollar</td>
</tr>
<tr>
<td>BIA</td>
<td>Budget impact analysis</td>
</tr>
<tr>
<td>CAD</td>
<td>Canadian Dollar</td>
</tr>
<tr>
<td>CEA</td>
<td>Cost-effectiveness analysis</td>
</tr>
<tr>
<td>CHEC</td>
<td>Consensus on Health Economic Criteria</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence interval</td>
</tr>
<tr>
<td>CT</td>
<td>Computed tomography</td>
</tr>
<tr>
<td>CUA</td>
<td>Cost-utility analysis</td>
</tr>
<tr>
<td>CXR</td>
<td>Chest X-ray</td>
</tr>
<tr>
<td>ELCAP</td>
<td>Early Lung Cancer Action Project</td>
</tr>
<tr>
<td>EUR</td>
<td>Euro</td>
</tr>
<tr>
<td>GBP</td>
<td>Great Britain Pound</td>
</tr>
<tr>
<td>GRADE</td>
<td>Grading of Recommendations Assessment, Development and Evaluation</td>
</tr>
<tr>
<td>HPFS</td>
<td>Health Professionals Follow-up Study</td>
</tr>
<tr>
<td>HTA</td>
<td>Health Technology Assessment</td>
</tr>
<tr>
<td>ICER</td>
<td>Incremental cost-effectiveness ratio</td>
</tr>
<tr>
<td>ICTRP</td>
<td>International Clinical Trial Registry Platform</td>
</tr>
<tr>
<td>IQWiG</td>
<td>Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen</td>
</tr>
<tr>
<td>JPY</td>
<td>Japanese Yen</td>
</tr>
<tr>
<td>IV</td>
<td>Inverse-variance</td>
</tr>
<tr>
<td>LDCT</td>
<td>Low-dose computed tomography</td>
</tr>
<tr>
<td>Lung-RADS</td>
<td>Lung Imaging Reporting and Data System</td>
</tr>
<tr>
<td>LUSI</td>
<td>Lung tumor screening and intervention trial</td>
</tr>
<tr>
<td>LYG</td>
<td>Life-year gained</td>
</tr>
<tr>
<td>M-H</td>
<td>Mantel-Haenszel method</td>
</tr>
<tr>
<td>MISCAN</td>
<td>Microsimulation Screening ANalysis-Lung</td>
</tr>
<tr>
<td>NHS</td>
<td>UK National Health Services</td>
</tr>
<tr>
<td>NHS EED</td>
<td>Economic Evaluation Database from the UK National Health Service</td>
</tr>
<tr>
<td>NLST</td>
<td>US National Lung Screening Trial</td>
</tr>
<tr>
<td>NIHR</td>
<td>UK National Institute for Health Research</td>
</tr>
<tr>
<td>OIS</td>
<td>Optimal information size</td>
</tr>
<tr>
<td>PanCan</td>
<td>Pan-Canadian Early Detection of Lung Cancer</td>
</tr>
<tr>
<td>PICO</td>
<td>Population, intervention, comparator and outcomes</td>
</tr>
<tr>
<td>PLCO</td>
<td>Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial</td>
</tr>
<tr>
<td>QALY</td>
<td>Quality-adjusted life year</td>
</tr>
<tr>
<td>RCT</td>
<td>Randomized controlled trial</td>
</tr>
<tr>
<td>RR</td>
<td>Relative risk</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>SEER</td>
<td>Surveillance Epidemiology and End Results</td>
</tr>
<tr>
<td>SMD</td>
<td>Standardized mean difference</td>
</tr>
<tr>
<td>UKLS</td>
<td>UK Lung Cancer Screening Trial</td>
</tr>
<tr>
<td>USD</td>
<td>United States Dollar</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
</tbody>
</table>
1 Background

Each year, lung cancer is responsible for 1.6 to 1.8 Mio death worldwide\textsuperscript{1,2} and for about 3200 death in Switzerland.\textsuperscript{3} Over 20\% of cancer-related death in Switzerland are caused by lung cancer.\textsuperscript{3} 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\%.\textsuperscript{4} 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-years 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. At least seven smaller trials, but still 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 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.\textsuperscript{5}

2 Aims of the Health Technology Assessment report

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.
3 Clinical effectiveness

3.1 Existing literature
A systematic search for evidence syntheses was conducted on January 20, 2020, in MEDLINE via Pubmed and the Cochrane library. The search strategy consisted of terms related to lung cancer and were combined with terms related to screening and LDCT. A total of 125 hits were screened. Thereof, 51 hits were published before 2014, including the only Cochrane review which was published in 2013, and were checked for their relevance, but will not be further considered in the present scope. Of the 74 hits published in 2014 or later, 56 were not relevant (no systematic evidence syntheses n=18, wrong population n=23, wrong intervention n=11, health economic analyses n=2, protocol n=1 and Chinese language n=1). Of the remaining 18 relevant evidence syntheses, seven included only randomized controlled trials (RCTs) (Section 3.1.1) and 11 also included non-randomized trials (Section 3.1.2).

3.1.1 Evidence syntheses on randomized controlled trials
The most recent Cochrane review on LDCT lung cancer screening was published in 2013. Trials comparing LDCT with no screening were still ongoing at this time point and no data on mortality was published. The most recent and high-quality evidence synthesis is the Health Technology Assessment report published in 2018 commissioned by the UK National Institute for Health Research (NIHR). The findings of this report were partly published in a peer-reviewed journal. The NIHR report provides a comprehensive overview of eight partly published and one ongoing randomized trials comparing LDCT with no screening. However, data on mortality was available for only three trials as the systematic literature search was conducted in January 2017. The pooled effect for lung cancer mortality of these three trials was not statistically significant (relative risk [RR] 1.15, 95% CI 0.79 to 1.67). A recently published systematic review Huang 2019 on LDCT lung cancer screening pooled lung cancer mortality data of seven trials and reported a RR of 0.78 (95% CI 0.68 to 0.89, I²=0%) in favor of the LDCT when compared to no screening. The potential reduction of lung cancer mortality is still subject to uncertainty because Huang 2019 pooled preliminary results of two RCTs (Yang 2018 and NELSON). Moreover, Huang 2019 did not analyze the potential harm from lung cancer screening due to diagnostic work-up and over-diagnosis. Six other evidence syntheses on RCTs were identified, two systematic reviews reported findings (including relevant RCTs) which were largely overlapping with the Huang 2019 or the NIHR report, and two systematic reviews were outdated.

3.1.2 Evidence syntheses also including non-randomized studies
Three recent evidence syntheses have considered randomized and non-randomized controlled trials. Interestingly, the number of participants from RCTs outnumbers the one from participants in non-randomized trials and so, most of the available evidence on harm derives from randomized controlled trials. For example, Usman 2016 concluded that a invasive procedures were part of follow-up investigations in 403 participants with benign conditions from the total of 40,569 participants, of whom 36,000 did participate in RCTs. Several available RCTs have a fairly large sample size of several thousand participants, whereas two RCTs form the largest body of evidence in term of participants, the NLST from the US and the NELSON trial from the Netherlands.

Several evidence syntheses assessed endpoints which were not relevant for the present scope: One systematic review assessed demographic differences in screening programs, one assessed only the
early cancer detection rates\textsuperscript{23}, two assessed how lung cancer screening influenced smoking behaviour\textsuperscript{24,25} and two investigated psychological burden of lung cancer screening.\textsuperscript{26,27}

Two systematic reviews\textsuperscript{28,29} evaluated lung cancer screening in an asbestos-exposed population. No RCTs were included. In the included cohorts, the number of events (mortality) were mostly very low or not reported, and long-term follow-up results were rare.

### 3.1.3 NLST - The National Lung Cancer Screening Trial

The National Lung Cancer Screening Trial randomized 53,454 participants with high risk for lung cancer to either LDCT (n=26,722) or chest radiography (n=26,732) in 33 US medical centers. Compared to chest radiography screening, LDCT screening reduced lung cancer mortality by 20\% (relative risk reduction: 20\%, 95\% CI 6.8 to 26.7).\textsuperscript{5,30} Because of its high internal validity and the large cohort size, NLST is considered as strongest evidence for the benefit of lung cancer screening. However, the beneficial effect on lung cancer mortality is accompanied by a considerable false-positive rate of 96.4\% in the LDCT and 94.5\% in the chest radiography group. The consequences of false-positive scans are expressed in absolute numbers for the LDCT group: 18,146 of all 75,126 CT scans over three screening rounds were positive. Of all positive scans 17,497 were false-positive, and most of them were subjected to follow-up investigations. Over 39\% of the participants in the LDCT group had at least one positive scan during the three screening rounds. In the NLST trial, all non-calified nodules with larger diameter of 4 mm were considered as positive scan, this relatively unspecific definition was probably the main driver for the high false-positive rates.

Importantly, the generalizability of the NLST for the Swiss setting is considered to be limited. First, Swiss institutions follow guidelines with a more complex definition for positive scans (e.g. following the Fleischner Society Guidelines\textsuperscript{31}), which results in substantial fewer false-positive results (see also below the definition of positive scans in the NELSON trial). Second, chest radiography is not recommended, mainly because no health benefit has been demonstrated from routine chest radiography screening in smoking individuals.\textsuperscript{32} Hence, NLST lacks a relevant comparator (e.g. no screening). Despite of these two important limitations for the Swiss setting, a network meta-analysis for the outcomes lung cancer mortality and all-cause mortality between LDCT, chest X-ray and no screening will be performed.

### 3.1.4 NELSON - Dutch-Belgian lung cancer screening trial

The Dutch-Belgian lung cancer screening trial (NELSON) is the largest trial comparing LDCT screening to no screening. NELSON randomized 7915 smokers (or former smokers) to LDCT screening and 7155 to no screening.\textsuperscript{33} NELSON was completed in 2018 and the recently published 10 years follow-up data showed that LDCT screening compared to no screening reduces lung cancer mortality in males (cumulative rate ration: 0.76, 95\% CI 0.61 to 0.94).\textsuperscript{34} The screening, however, did not affect the all-cause mortality. The NELSON trial reported a false-positive scans rate of 56\%\textsuperscript{33,34}, and hence, much lower than in the NLST. In absolute numbers, 467 (2.1\%) of the 22,600 scans during three screening rounds were positive scans, thereof, 264 were false-positive scans. In the NELSON trial, the definition of a positive scan was based on the volume of the nodule or its growth rate. Scans were classified into negative, positive or indeterminate. Indeterminate scans required follow-up LDCT and were then classified into positive or negative scan based on lesion volume doubling time.\textsuperscript{35}
3.1.5 Ongoing trials

The Clinicaltrials.gov register was systematically searched for ongoing RCTs on December 4, 2019. Ninety-one trials were screened, and three ongoing trials were identified, one Canadian and two Chinese trials (Table 1). The one Chinese trial (NCT03975504) is a follow-up trial of the other Chinese trial (NCT02898441).

Table 1: Ongoing trials

<table>
<thead>
<tr>
<th>Clinicaltrials.gov Country</th>
<th>Planned start Completion date</th>
<th>N participants to be enrolled</th>
<th>Published results</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCT0243196236 Canada</td>
<td>April, 2015 December 2019</td>
<td>800 at risk</td>
<td>Only on smoking cessation rates37</td>
</tr>
<tr>
<td>NCT0289844138 China</td>
<td>January 1, 2014 December 2018</td>
<td>6000 participants with high-risk for lung cancer</td>
<td>Screening observation after first round, mortality rates for 2 year follow-up13</td>
</tr>
<tr>
<td>NCT0397550439 China</td>
<td>August 1, 2018 July 31, 2023</td>
<td>6000 high-risk subjects</td>
<td>No</td>
</tr>
</tbody>
</table>

3.2 Aim

The assessment of the clinical effectiveness aims to summarize the highest available evidence to inform health policy decision-makers about the potential benefit and harm of lung cancer screening with LDCT in smokers and former smokers compared to no screening.

Following the submission of the scoping in May 2020 a preliminary HTA report was published by IQWiG on June 30, 2020 covering the same topic (Gesundheitswesen, 2020 #259). The purchaser then requested to the revise the scope and to provide solely an updated report based on the HTA report from the UK by Snowsill et al. (Snowsill, 2018 #12)

In addition to the direct comparison of LDCT screening with no screening, a network meta-analysis will be performed taking into account RCTs comparing LDCT screening with chest X-ray (e.g. NLST) and RCTs comparing chest X-ray with no screening. This “triangular” network of three screening strategies will be limited to the outcomes lung cancer mortality and all-cause mortality. Details for the network meta-analysis are described in Section 3.3.9.

3.3 Methods

3.3.1 Overview of the eligibility criteria

The overview of eligibility criteria (PICO-Question) used in the literature selection process is shown in Table 2.

Table 2: PICO-Question for the assessment of clinical effectiveness
### 3.3.2 Eligibility criteria

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

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

#### 3.3.2.3 Comparator
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 3.3.9).

#### 3.3.2.4 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 → A indeterminate scan is defined as a scan which does not allow to classify 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 assessment 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.

3.3.2.5 Study design
Relevant study designs will include 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 randomization are deemed inadequate, these study types will be considered because it can be assumed that individuals in such studies were prospectively assigned to the intervention or the comparator.\(^{40}\)

3.3.2.6 Languages
Trials published in English, French, and German will be eligible for inclusion.

3.3.3 Literature search
The literature search will comprise Medline ALL and EMBASE via OvidSP, CINAHL (“Cumulative Index to Nursing and Allied Health Literature”) via EBSCO and CENTRAL (“Cochrane central register of controlled trials”). In addition, reference lists of systematic reviews will be screened for trials that fulfill the inclusion criteria. The same search strategy will be applied and updated as in the HTA report by Snowsill et al. which were done in December 2016 and January 2017. \{Snowsill, 2018 #12\}

The topic-specific search strategy will be combined with a search filter for randomized controlled trials (RCTs). The search strategy will not be restricted by adding terms for the comparator. Conference proceedings or conference booklets will not be searched.

Two reviewers will independently screen titles/abstracts of records found in the literature search for potentially eligible studies after removal of duplicate publications. Subsequently, two reviewers independently will screen the full-text articles of the potentially eligible studies in order to identify eligible RCTs. Discrepant screening results will be discussed and will be resolved by consensus or by third-party arbitration. Protocols of included RCTs will be searched for within the US trial registry (clinicaltrials.gov) and WHO trial registry.

The US trial registry (clinicaltrials.gov) and WHO trial registry (International Clinical Trial Registry Platform, ICTRP) will be searched for unpublished or still ongoing trials.

3.3.4 Data extraction
Data on study characteristics and patient-relevant outcomes of trials identified and included in the report by Snowsill et al. will be re-extracted into a standardized form by one reviewer, and updated
with newly published trials and checked by another. Discrepancies will be resolved by discussion or third-party arbitration.

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) will be extracted. General study population characteristics (age, sex, smoking behavior/status, etc.) and characteristics of the lung cancer-positive population (cancer stage, histologic type, etc.) will be extracted. Radiation exposure will not be extracted, but will be discussed in the HTA report and existing literature on radiation exposure will be referenced.

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

Continuous outcome data will 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 will be extracted and analyzed, and not the number of events themselves. If only the number of events will be available, this information will be extracted and will be summarized in the relevant sections. Pooling of number of events will only be considered if consistently reported by all trials.

For missing information, study authors will not be contacted.

### 3.3.5 Risk of bias and quality of evidence assessment

One reviewer will assess the internal validity (risk of bias assessment) of each trial. This will be checked by a second reviewer. Discrepancies will be resolved by discussion or third-party arbitration.

To assess the risk of bias of individual trials the following criteria will be used:

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

- The risk of attrition bias will be judged low if the proportion of patients with missing data is 0 - 10% in either study arm and comparable between the randomized treatment arms.
- The risk of attrition bias will also be judged low if the proportion of patients with missing data is between 10-20% per arm, is comparable between the randomized treatment arms, and is 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 will be considered comparable if the difference between the intervention and control group are 5% or less.
- The risk of attrition bias will be judged high if more than 20% of the data were missing irrespective of how the missing data were addressed in the analysis.

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

The quality of the evidence will be judged by one reviewer 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 will be resolved by consensus or third-party arbitration. The following criteria will be considered to judge the quality of the evidence41-56:

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 continuous outcomes, the precision will be adequate if the optimal information size (OIS) is sufficient (simple sample size calculation to estimate whether the total number of included patients would be sufficient for an adequately powered RCT) and for binary outcomes, if the number of events is sufficient (rule of thumb >300 events).47 If the sample size or number of events is sufficiently large, the 95% CI of the effect estimate will be examined. If the 95% CI is narrow enough not to include both the “no effect” line and a possible clinically relevant effect (also called minimal clinically important difference) precision will be judged as adequate.47

Using the GRADEpro GDT software57 results of the judgment will be presented in a summary of findings table.

3.3.6 Data synthesis
Study characteristics and results of the eligible trials will be presented per study in tables and will be descriptively summarised.
Where possible, outcome results will be summarised quantitatively in a meta-analysis by using a random-effects model. Therefore, the inverse-variance (IV) method\textsuperscript{58} for continuous outcomes and the Mantel-Haenszel method\textsuperscript{40} (M-H) for binary outcomes will be applied.

Continuous outcomes will be presented as mean differences. For binary outcomes, relative risks (RR) will be determined. Effect estimates (summary and single for each trial) with the corresponding 95% confidence interval will be presented in forest plots.

If a continuous outcome is measured on different scales, mean differences of the individual trial results will be standardized using the following formula:

$$\text{Standardized mean difference (SMD)} = \frac{\text{mean}_{\text{intervention}} - \text{mean}_{\text{comparator}}}{\text{SD}_{\text{pooled}}}$$

An effect size above 0.2 SDs will be considered to correspond to a small effect; effect sizes above 0.5 SDs to a medium effect and above 0.8 SDs will be considered to correspond to large effects\textsuperscript{59,60}.

Heterogeneity of pooled effect estimates will be estimated using $I^2$. Estimates of $I^2$ will be interpreted under the guidance of the Cochrane Handbook\textsuperscript{40}. Heterogeneity with an $I^2$ of 0% to 40% will be considered low, 41% to 60% will be considered moderate, and 61% to 100% will be considered high. The interpretation of the observed $I^2$ value will depend on other measures for heterogeneity, namely Tau\textsuperscript{2} (a Tau\textsuperscript{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\textsuperscript{40,61}.

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

### 3.3.7 Subgroup analyses

To assess possible variations of treatment effects the following subgroup analyses will be considered:

- 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 will be assessed by interaction tests available within Review Manager 5.3 and according to the Cochrane Handbook.\textsuperscript{40}

### 3.3.8 Sensitivity analyses

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

A network meta-analysis will be performed in addition to the direct comparison of LDCT screening with no screening. The random-effects network meta-analysis will be performed for two outcomes lung cancer mortality and all-cause mortality, and consists of three connected nodes (LDCT, chest X-ray and usual care/no screening) (Figure 1). Based on the data availability, either a Bayesian or frequentist random-effects method will be used. 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 will be reported as RRs with corresponding credibility intervals.\(^{62-64}\) Statistical analysis will be performed using an R package “gemtc”. The confidence in the results of the network meta-analysis will be assessed with CINeMA.\(^{65}\)

![Figure 1 Network map of the network meta-analysis](image)
4 Health economic evaluation

Treatment for late-stage lung cancer is rarely curative and expensive. Screening people at high risk of lung cancer before symptomatic cancer develops may help to treat them more effectively. The costs related to a screening program, including the consultation, the diagnostic, and the treatment costs (also related to false-positive results and adverse events) may be considerably high, especially if the prevalence of subjects at high risk is elevated. Investigating the cost-effectiveness and budget impact of different screening options (e.g. subject inclusion criteria, screening intensity) is therefore fundamental to identify the screening scenarios with the best balance between benefits, harms, and costs.

4.1 Aim

The health economic assessment aims to investigate the cost-effectiveness and the potential budget impact of LDCT screening programs for lung cancer in high-risk populations (i.e. smokers and ex-smokers) in Switzerland.

4.2 Pre-review of the health economic literature

As part of the scoping process, a preliminary search for health economic literature was conducted in PubMed (MEDLINE) to gain a first understanding of potentially relevant cost-effectiveness studies on lung cancer screening (e.g. cost-effectiveness, cost-utility, or cost studies). The following search strings were combined:

- Lung cancer (n=336,415)
- Screening (n=4,647,558)
- LDCT OR low-dose CT OR low-dose computed tomography OR CT OR computed tomography (n=769,110)
- Afford$ OR Budget$ OR Capital expenditure$ OR cost$ OR cost-benefit OR Cost-consequence$ OR Cost-effectiveness OR Cost-minimization OR Cost-utility OR Economic$ OR Economic evaluation OR Expenditure$ OR Fee$ OR Finance$ OR Financial OR Financing OR Health expenditure$ OR Health resource allocation OR Health resource utilization OR Health economic$ OR Medical savings accounts OR Monetary OR Pharmaco-economic analysis OR Pharmaco-economic analysis OR Pharmacoeconomic$ OR Pharmacoeconomic analyses OR Pharmacoeconomic analysis OR Price$ OR Socioeconomic$ (n=1,569,412)

The search conducted on 18 December 2019 resulted in 1,240 hits. A first title/abstract screen led to the identification of 39 potentially relevant articles published between 2001 and 2019. Among them, there were four HTAs, 30 cost-effectiveness or cost-utility analyses, three systematic reviews, and two narrative reviews.

After full-text review, ten articles were excluded for following reasons:

- Outdated Swedish HTA (published in 2003, i.e. before most of the trials on this topic were published)
Non-systematic reviews (CADTH 2015, Lew 2019)\textsuperscript{66,67}
- Outdated systematic reviews (Puggina 2016, Raymakers 2016)\textsuperscript{68,69}
- Systematic review in Chinese (Liu 2019)\textsuperscript{70}
- Wrong intervention/comparators (Allen 2019, Hinde 2018, Kumar 2018)\textsuperscript{71-73}
- Unclear population (i.e. no information on smoking behaviour of selected patients) (Kanarkievitz 2015)\textsuperscript{74}

4.3 Brief overview of the identified cost-effectiveness analyses
The characteristics of the identified articles are summarized in Table 3.

Two-thirds of the identified articles were included in the most recent HTA conducted by Snowsill et al. on behalf of the National Institute for Health Research (NIHR) in the UK.\textsuperscript{7} Snowsill et al. did not include a Canadian HTA published in 2014 and another eight studies published between 2017 and 2018.

All cost-effectiveness analyses included high risk population for lung cancer (mainly using the NLST inclusion criteria). The interventions ranged from one single LDCT screen to annual, biennial, or triennial screens from patient inclusion until 80 years of age. The reported comparator was no screening. However, it is important to remark that many recent publications used the results of the chest X-ray control arm of the NLST as comparator, assuming that chest X-ray is equivalent to no screening by assuming that annual screening with chest radiograph does not reduce lung cancer mortality compared with usual care.\textsuperscript{75} While this assumption might hold for the endpoint of overall mortality this might not be true for other endpoints and in particular in regard to costs for work-up of false positive x-ray results.

The main results of the identified articles are summarized in Table 4. Similar to HTA report by Snowsill et al., the existing economic evaluations of LDCT screening for lung cancer have produced markedly variable estimates of the cost-effectiveness of screening.\textsuperscript{7} The identified studies, using different methodological approaches as well as different effectiveness sources, reported incremental cost-effectiveness ratios (ICERs) for LDCT screening ranging from few thousands of USD to more than USD 100,000 per quality-adjusted life year (QALY) gained. Assuming a cost-effectiveness threshold of USD 100,000 per QALY, most of the identified cost-effectiveness analyses would suggest that LDCT is cost-effective if compared to no screening.

Among the identified study, one was investigating the cost-effectiveness of LDCT screening for lung cancer in Switzerland.\textsuperscript{76} This study adapted the Mlcrosimulation Screening ANalysis-Lung (MISCAN) model and used Swiss-specific input parameters for lung cancer epidemiology, smoking behaviour, and treatment costs. The model parameters for the effectiveness of CT screening were calibrated to individual-level data from the NLST. The study concluded that several LDCT screening strategies may be cost-effective in Switzerland (i.e. would show an incremental cost-effectiveness ratio below EUR 50,000 per life-year gained).
<table>
<thead>
<tr>
<th>Study author and publication year</th>
<th>Type of evaluation</th>
<th>Location, price year and currency</th>
<th>Population</th>
<th>Intervention(s)</th>
<th>Comparator(s)</th>
<th>Methodology</th>
</tr>
</thead>
<tbody>
<tr>
<td>HTA Ontario 2014</td>
<td>HTA (CUA)</td>
<td>Alberta, 2012 CAD</td>
<td>NLST cohort (aged 55–74 years with ≥ 30 pack-year smoking history)</td>
<td>Annual and biennial LDCT screening</td>
<td>No screening</td>
<td>Microsimulation model</td>
</tr>
<tr>
<td>HTA Field 2016</td>
<td>HTA (CEA and CUA)</td>
<td>UK, 2011–12 GBP</td>
<td>Adults aged 50–75 years</td>
<td>Risk prediction followed by single LDCT screen</td>
<td>No screening</td>
<td>Decision tree model</td>
</tr>
<tr>
<td>HTA Snowsill 2018</td>
<td>HTA</td>
<td>UK, 2016 GBP</td>
<td>Adult smokers (current or former) aged 55–80</td>
<td>Single, triple, annual, or biennial LDCT screen</td>
<td>No screening</td>
<td>Decision tree model</td>
</tr>
<tr>
<td>Black 2014</td>
<td>CEA and CUA</td>
<td>USA, 2009 USD</td>
<td>NLST cohort (aged 55–74 years with ≥ 30 pack-year smoking history)</td>
<td>Annual LDCT for 3 years, Annual CXR for 3 years</td>
<td>No screening</td>
<td>Decision tree model</td>
</tr>
<tr>
<td>Black 2015</td>
<td>CEA and CUA</td>
<td>USA, 2009 USD</td>
<td>NLST cohort (aged 55–74 years with ≥ 30 pack-year smoking history)</td>
<td>Annual LDCT for 3 years, Annual CXR for 3 years</td>
<td>No screening</td>
<td>Decision tree model</td>
</tr>
<tr>
<td>Chirikos 2003</td>
<td>CEA</td>
<td>USA, 2000 USD</td>
<td>Adult smokers aged 45–74 years</td>
<td>Annual LDCT for 5 years</td>
<td>No screening</td>
<td>Cohort model</td>
</tr>
<tr>
<td>Cressmann 2017</td>
<td>CUA</td>
<td>Canada, 2015 CAD</td>
<td>NLST cohort (aged 55–74 years with ≥ 30 pack-year smoking history)</td>
<td>Annual LDCT screening</td>
<td>CXR (assumed to be equal to no screening)</td>
<td>Decision tree model</td>
</tr>
<tr>
<td>Criss 2019</td>
<td>CEA</td>
<td>USA, 2018 USD</td>
<td>Current, former, and never-smokers aged 45 years from the 1960 U.S. birth cohort</td>
<td>Annual LDCT according to NLST, CMS, and USPSTF criteria.</td>
<td>No screening</td>
<td>Microsimulation models</td>
</tr>
<tr>
<td>Goffin 2015</td>
<td>CUA</td>
<td>Canada, 2008 CAD</td>
<td>Smokers aged 55–74 years with ≥ 30 pack-year history</td>
<td>Annual LDCT for 3 years with smoking cessation</td>
<td>No intervention</td>
<td>Microsimulation model</td>
</tr>
<tr>
<td>Goffin 2016</td>
<td>CUA</td>
<td>Canada, 2008 CAD</td>
<td>Smokers aged 55–74 years with ≥ 30 pack-year history</td>
<td>Biennial LDCT screening for 20 years with/without smoking cessation</td>
<td>Annual LDCT screening for 20 years with/without smoking cessation</td>
<td>Microsimulation model</td>
</tr>
<tr>
<td>Study</td>
<td>Year</td>
<td>Type</td>
<td>Country/Currency</td>
<td>Eligibility</td>
<td>Screening Frequency</td>
<td>Intervention</td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td>------</td>
<td>------------------</td>
<td>-------------</td>
<td>---------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Goulart 2012</td>
<td>CEA</td>
<td>USA, 2011 USD</td>
<td>Those eligible for NLST, i.e. smokers aged 55 to 74 years</td>
<td>LDCT screening (frequency unclear)</td>
<td>No screening</td>
<td>Decision tree model</td>
</tr>
<tr>
<td>Hofer 2018</td>
<td>CUA</td>
<td>Germany, 2016 EUR</td>
<td>Ever smokers aged 55–75 with ≥ 20 cigarettes per day</td>
<td>Annual LDCT screening</td>
<td>No screening</td>
<td>Markov model</td>
</tr>
<tr>
<td>Jaie 2018</td>
<td>CUA</td>
<td>New Zealand, 2011 USD</td>
<td>NLST cohort (aged 55–74 years with ≥ 30 pack-year smoking history)</td>
<td>Biennial LDCT screening</td>
<td>CXR (assumed to be equal to no screening)</td>
<td>Microsimulation model</td>
</tr>
<tr>
<td>Mahadevia 2003</td>
<td>CEA</td>
<td>USA, 2001 USD</td>
<td>60-year-old heavy smokers (current and former, &gt; 20 pack per day)</td>
<td>Annual LDCT to age 80 years</td>
<td>No screening</td>
<td>Markov model</td>
</tr>
<tr>
<td>Manser 2004</td>
<td>CEA and CUA</td>
<td>Australia, 2002 AUD</td>
<td>Male current smokers aged 60–64 years</td>
<td>Annual LDCT for 5 years</td>
<td>No screening</td>
<td>Markov model</td>
</tr>
<tr>
<td>Marshall 2001</td>
<td>CEA</td>
<td>USA, 1999 USD</td>
<td>General smokers aged 60–74 years</td>
<td>Single LDCT screen</td>
<td>No screening</td>
<td>Decision tree model</td>
</tr>
<tr>
<td>McMahon 2011</td>
<td>CUA</td>
<td>USA, 2006 USD</td>
<td>Current and former smokers ≥20 pack-years smoking history</td>
<td>Annual LDCT screening</td>
<td>No intervention</td>
<td>Patient-level microsimulation model</td>
</tr>
<tr>
<td>Pyenson 2012</td>
<td>CEA</td>
<td>USA, 2012 USD</td>
<td>Current and former smokers aged 50 years with ≥ 30 pack-year smoking history</td>
<td>Annual LDCT from age 50–64 years</td>
<td>No screening</td>
<td>Cohort model</td>
</tr>
<tr>
<td>Pyenson 2014</td>
<td>CEA</td>
<td>USA, 2014 USD</td>
<td>Adults aged 55–80 years with ≥ 30 pack-year smoking history</td>
<td>Annual LDCT from age 50–64 years</td>
<td>No screening</td>
<td>Cohort model</td>
</tr>
<tr>
<td>Shmueli 2013</td>
<td>CUA</td>
<td>Israel, 2011 USD</td>
<td>Adults aged ≥45 years with ≥10 pack-year smoking history</td>
<td>Single LDCT screen</td>
<td>No screening</td>
<td>Decision tree model</td>
</tr>
<tr>
<td>Tabata 2014</td>
<td>CEA</td>
<td>Japan, JPY (price year unclear)</td>
<td>Smokers aged 55–74 years</td>
<td>Annual LDCT</td>
<td>Annual CXR</td>
<td>Decision tree model</td>
</tr>
<tr>
<td>Ten Haaf 2017</td>
<td>CEA</td>
<td>Canada, 2015 CAD</td>
<td>Adult smokers (current or former) aged 46–75</td>
<td>Eligibility criteria and annual or biennial LDCT screening</td>
<td>No screening</td>
<td>Microsimulation model</td>
</tr>
<tr>
<td>Tomonaga 2018</td>
<td>CEA</td>
<td>Switzerland, 2015 EUR</td>
<td>Adult smokers (current or former) aged 50–80</td>
<td>Eligibility criteria and annual, biennial, or triennial LDCT screening</td>
<td>No screening</td>
<td>Microsimulation model</td>
</tr>
<tr>
<td>Treskova 2017</td>
<td>CEA</td>
<td>Germany, EUR (price)</td>
<td>Smokers aged 55–74 years with ≥ 30 pack-year history</td>
<td>Annual LDCT for 5 years</td>
<td>No screening</td>
<td>Microsimulation model</td>
</tr>
<tr>
<td>Study</td>
<td>Year and Location</td>
<td>Target Population</td>
<td>Screening Strategy</td>
<td>Follow-up</td>
<td>Modeling Approach</td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>-------------------</td>
<td>-------------------</td>
<td>--------------------</td>
<td>-----------</td>
<td>-------------------</td>
<td></td>
</tr>
<tr>
<td>Villanti 2013</td>
<td>USA, 2012 USD</td>
<td>Adults aged 50-64 years with ≥ 30 pack-year history</td>
<td>Annual LDCT screening to age 64 years</td>
<td>No screening</td>
<td>Cohort model</td>
<td></td>
</tr>
<tr>
<td>Wade 2018</td>
<td>Australia, 2015 AUD</td>
<td>Smokers aged 55–74 years with ≥ 30 pack-year history</td>
<td>Annual LDCT for 3 years</td>
<td>CXR (assumed to be equal to no screening)</td>
<td>Markov model (not clearly stated)</td>
<td></td>
</tr>
<tr>
<td>Whynes 2008</td>
<td>UK, 2004 GBP</td>
<td>Men aged 61 years at high risk</td>
<td>Single LDCT screen</td>
<td>No screening</td>
<td>Decision tree model</td>
<td></td>
</tr>
<tr>
<td>Wisnivesky 2003</td>
<td>USA, 2000 USD</td>
<td>Adults aged ≥60 years with ≥10 pack-year smoking history</td>
<td>Single LDCT screen</td>
<td>No screening</td>
<td>Decision tree model</td>
<td></td>
</tr>
<tr>
<td>Yang 2017</td>
<td>Taiwan, 2013 USD</td>
<td>NLST cohort (aged 55–74 years with ≥ 30 pack-year smoking history)</td>
<td>Annual LDCT for 3 years</td>
<td>No screening</td>
<td>Markov model (not clearly stated)</td>
<td></td>
</tr>
</tbody>
</table>

* Study included in the HTA published by Snowsill et al. in 2018.

CEA: cost-effectiveness analysis; CUA: cost-utility analysis; CXR: chest x-ray; HTA: health technology assessment; LDCT: low-dose computed tomography
<table>
<thead>
<tr>
<th>Study author and publication year</th>
<th>Source of effectiveness</th>
<th>Health outcomes</th>
<th>Time horizon and discount rate</th>
<th>Main results</th>
</tr>
</thead>
<tbody>
<tr>
<td>HTA Ontario 2014</td>
<td>NLST</td>
<td>QALYs</td>
<td>Lifetime, 3%</td>
<td>LDCT screening more costly and more effective than no screening, ICER CAD 92,025/QALY (annual) or CAD 67,396 (biennial)</td>
</tr>
<tr>
<td>HTA Field 2016 #</td>
<td>UKLS and estimates of lead time</td>
<td>Life-years, QALYs</td>
<td>Lifetime, 3.5%</td>
<td>LDCT screening more costly and more effective than no screening, ICER GBP 8,466/QALY</td>
</tr>
<tr>
<td>HTA Snowsill 2018 #</td>
<td>Natural history model calibrated to UKLS and NLST</td>
<td>QALYs</td>
<td>Lifetime, 3.5%</td>
<td>LDCT screening more costly and more effective than no screening, ICER GBP 28,169-40,034/QALY</td>
</tr>
<tr>
<td>Black 2014 #</td>
<td>NLST (assume same outcomes for no screening as CXR)</td>
<td>Life-years, QALYs</td>
<td>Lifetime, 3%</td>
<td>LDCT screening more costly and more effective than CXR and no screening, ICER (vs. no screening) USD 81,000/QALY. CXR dominated by no screening</td>
</tr>
<tr>
<td>Black 2015 #</td>
<td>Hypothetical stage shift</td>
<td>Life-years</td>
<td>15 years, 7.5%</td>
<td>LDCT more expensive and more effective than no screening, ICER USD 33,557–90,022/LYG depending on achieved stage distribution</td>
</tr>
<tr>
<td>Cressmann 2017</td>
<td>NLST and PanCan</td>
<td>QALYs</td>
<td>Lifetime, 3%</td>
<td>LDCT screening more costly and more effective than no screening, ICER CAD 20,724/QALY</td>
</tr>
<tr>
<td>Criss 2019</td>
<td>NHS/HPFS, SEER, NLST, PLCO, Lung-RADS</td>
<td>Life-years, QALYs</td>
<td>45 years, 3%</td>
<td>LDCT screening more costly and more effective than no screening, ICER USD 36,000-51,900/LYG or USD 49,200-96,700/QALY</td>
</tr>
<tr>
<td>Goffin 2015 #</td>
<td>Natural history model, partially calibrated to NLST</td>
<td>QALYs</td>
<td>20 years (lifetime), 3%</td>
<td>LDCT screening more costly 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)</td>
</tr>
<tr>
<td>Goffin 2016 #</td>
<td>Natural history model partially calibrated to NLST</td>
<td>QALYs</td>
<td>Lifetime, 3%</td>
<td>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</td>
</tr>
<tr>
<td>Study</td>
<td>Model</td>
<td>Setting</td>
<td>Time horizon</td>
<td>Discounting</td>
</tr>
<tr>
<td>-------</td>
<td>-------</td>
<td>---------</td>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Goulart 2012 #</td>
<td>NLST</td>
<td>Lung cancer deaths</td>
<td>Unclear (possibly 1 year), no discounting</td>
<td>LDCT more expensive and more effective than no screening, ICER USD 240,000 per lung cancer death avoided</td>
</tr>
<tr>
<td>Hofer 2018</td>
<td>NLST, German LUSI trial</td>
<td>Life-years, QALYs</td>
<td>15 years, 3%</td>
<td>LDCT screening more costly and more effective than no screening, ICER EUR 19,302/LYG or EUR 30,291/QALY</td>
</tr>
<tr>
<td>Jaine 2018</td>
<td>NLST</td>
<td>QALYs</td>
<td>Lifetime, 3%</td>
<td>LDCT more expensive and more effective than no screening, ICER USD 104,000/QALY</td>
</tr>
<tr>
<td>Mahadevia 2003 #</td>
<td>Hypothetical stage shift</td>
<td>QALYs</td>
<td>40 years (to age 100), 3%</td>
<td>LDCT more expensive and more effective than no screening, ICER USD 116,300/QALY</td>
</tr>
<tr>
<td>Manser 2004 #</td>
<td>Diagnostic performance of LDCT based on ‘weighted averages of six studies’</td>
<td>Life-years, QALYs</td>
<td>15 years, 3%</td>
<td>LDCT more expensive and more effective than no screening, ICER AUD 57,325/LYG or AUD 105,090/QALY</td>
</tr>
<tr>
<td>Marshall 2001 #</td>
<td>ELCAP</td>
<td>Life-years</td>
<td>5 years, 3%</td>
<td>LDCT more expensive and more effective than no screening, ICER USD 23,100/LYG. In ‘very high-risk’ cohort, ICER USD 5,940/LYG</td>
</tr>
<tr>
<td>McMahon 2011 #</td>
<td>Natural history model calibrated to tumour registry data and validated against screening studies</td>
<td>QALYs</td>
<td>Lifetime, 3%</td>
<td>LDCT more expensive and more effective than no screening. ICERs for screening consistently above USD100,000/QALY unless positive impact on smoking cessation included</td>
</tr>
<tr>
<td>Pyenson 2012 #</td>
<td>ELCAP</td>
<td>Life-years</td>
<td>15 years, no discounting</td>
<td>LDCT more expensive and more effective than no screening, ICER USD 18,862/LYG</td>
</tr>
<tr>
<td>Pyenson 2014 #</td>
<td>ELCAP</td>
<td>Life-years</td>
<td>20 years, no discounting</td>
<td>LDCT screening more costly and more effective than no screening, ICER USD 18,452/LYG</td>
</tr>
<tr>
<td>Shmueli 2013 #</td>
<td>Single-centre Israeli cohort study</td>
<td>QALYs</td>
<td>Lifetime, 3%</td>
<td>LDCT more expensive and more effective than no screening, ICER USD 1,464/QALY</td>
</tr>
<tr>
<td>Tabata 2014 #</td>
<td>ALCA, Japanese case–control study</td>
<td>Life-years</td>
<td>Unclear</td>
<td>LDCT screening more costly and more effective than CXR, ICERs ranging from JPY 983,000 to JPY 1,942,000/LYG depending on sex and age</td>
</tr>
<tr>
<td>Ten Haaf 2017 #</td>
<td>Natural history model calibrated to NLST</td>
<td>Life-years</td>
<td>Lifetime, 3%</td>
<td>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</td>
</tr>
<tr>
<td>Study</td>
<td>Model Calibration</td>
<td>Outcomes</td>
<td>Discounting</td>
<td>Comments</td>
</tr>
<tr>
<td>-------</td>
<td>-------------------</td>
<td>----------</td>
<td>-------------</td>
<td>----------</td>
</tr>
<tr>
<td>Tomonaga 2018</td>
<td>Natural history model calibrated to NLST</td>
<td>Life-years</td>
<td>Lifetime, 3%</td>
<td>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 &lt; EUR 50,000 per LYG.</td>
</tr>
<tr>
<td>Treskova 2017</td>
<td>Natural history model calibrated to tumour registry data and validated against screening studies</td>
<td>Life-years</td>
<td>Lifetime, 3%</td>
<td>LDCT screening more costly and more effective than no screening, ICER EUR 16,754-23,847/LYG</td>
</tr>
<tr>
<td>Villanti 2013 #</td>
<td>ELCAP and NLST</td>
<td>QALYs</td>
<td>15 years, no discounting</td>
<td>LDCT more expensive and more effective than no screening, ICER USD 28,240/QALY (ELCAP) or USD47,115/QALY (NLST). Adding smoking cessation nearly doubled QALY gain from screening alone and had lower ICER</td>
</tr>
<tr>
<td>Wade 2018</td>
<td>NLST</td>
<td>Life-years, QALYs</td>
<td>10 years, 5% discount</td>
<td>LDCT more expensive and more effective than no screening, ICER AUD 138,000/LYG or AUD 233,000/QALY</td>
</tr>
<tr>
<td>Whynes 2008 #</td>
<td>ELCAP</td>
<td>QALYs</td>
<td>Unclear (perhaps 40 years), 3.5%</td>
<td>LDCT more expensive and more effective than no screening, ICER GBP 13,910/QALY (for men)</td>
</tr>
<tr>
<td>Wisnivesky 2003 #</td>
<td>ELCAP</td>
<td>Life-years</td>
<td>Unclear, 3%</td>
<td>LDCT more expensive and more effective than no screening, ICER USD 2,500/LYG</td>
</tr>
<tr>
<td>Yang 2017</td>
<td>NLST</td>
<td>Lifetime, 3%</td>
<td>LDCT more expensive and more effective than no screening, ICER USD 19,683/QALY</td>
<td></td>
</tr>
</tbody>
</table>

# Study included in the HTA published by Snowsill et al. in 2018.

4.4 Approach to the health economic assessment

The health economic assessment aims to investigate the cost-effectiveness and the potential budget impact of LDCT screening programs for lung cancer in high-risk populations in Switzerland.

The assessment will include a systematic review of the currently published cost-effectiveness evidence (including an adaptation of the cost results to Switzerland), an update of the Swiss cost-effectiveness analysis published in 2018 (Tomonaga et al.76), and a budget impact analysis (BIA) based on Swiss data.

4.4.1 Systematic review of existing cost-effectiveness evidence

A systematic review of the current economic literature will be undertaken. The systematic review will be based on the HTA published by Snowsill et al. in 2018. 7 The aim will be 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 (the database search by Snowsill et al was conducted the 5th of January 2017).

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

4.4.1.1 Literature search strategy

The search strategy published by Snowsill et al. will be used to identify all relevant articles. 7 Following electronic databases will be used: Medline ALL, EMBASE, and the Cochrane Library and the Centre for Review and Dissemination (CRD) database including the Database of Abstracts of Reviews of Effects (DARE), Cochrane reviews, Health Technology Assessments (HTA) and the Economic Evaluation Database from the UK National Health Service (NHS EED). In Snowsill et al. articles identified in Medline (1,073 hits), EMBASE (1,314 hits), and Cochrane (49 hits) represented 84% of the identified articles (2,436 out of 2,915 hits including duplicates). Considering that the total of unique records, i.e. the total after removing all duplicates, was 2,223, the combined search in Medline, EMBASE, and Cochrane will presumably cover the great majority of published cost-effectiveness analyses and other relevant articles. Compared the search conducted by Snoswill et al., we will not perform a literature search in the Health Management Information Consortium (HMIC, 22 Hits in Snowsill et al.) in the Web of Science (358 hits), and in the EconLit (99 hits).

The literature search will be conducted by specialists of the Basel University Library, and the period of search will be limited to all articles published from 2017 onwards.

4.4.1.2 Screening of the search results

The results of the literature search will be screened by title, abstract and, if necessary, by full text review. In a first step, title and abstracts will be screened for relevant quantitative results (e.g. costs, LYG, QALYs, or ICERs) or for sentences suggesting potentially relevant content in the full text version.

Potentially relevant abstracts proceeded to the next step, in which full texts will be screened. Articles will then be classified in three groups:

- Relevant articles: full scale cost-effectiveness analyses using a PICO corresponding to the present scoping and reporting an endpoint of cost per QALY gained or cost per life-year gained. Ideally the analysis should be performed for a jurisdiction with broadly similar socioeconomic
characteristics as Switzerland (e.g. North, Central, and Western European countries, the USA, Canada, Australia, and New Zealand).

- Articles potentially providing important additional information: articles not meeting the criteria for the 'relevant' category but potentially containing useful additional information concerning effectiveness or costs, and thus being 'partially relevant'. Depending on the quality and quantity of information available from relevant articles, some partially relevant articles will be used as an additional source of information and for comparison.

- Irrelevant articles.

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

Methodological quality will be assessed using the Consensus on Health Economic Criteria (CHEC)-list for economic evaluations as in the HTA conducted by Snowsill et al. 7

International cost-effectiveness studies will be 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 et 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 will be 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 have to meet the 'PIC', or have to be performed for countries similar to Switzerland in terms of socioeconomic characteristics.

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

- CHEC 1: Is the study population clearly described?
- CHEC 2: Are competing alternatives clearly described? (intervention, comparator)
- CHEC 5: Is the chosen time horizon appropriate in order to include relevant costs and consequences?
- CHEC 6: Is the actual perspective chosen appropriate?
- CHEC 7: Are all important and relevant costs for each alternative identified?
- CHEC 9: Are costs valued appropriately? (currency, price date, conversion)
- CHEC 10: Are all important and relevant outcomes for each alternative identified?
- CHEC 13: Is an incremental analysis of costs and outcomes of alternatives performed?

The availability of costs and outcomes of interest for both the intervention and the comparator strategies will be considered fundamental.

4.4.1.4 Extraction of information
For the eligible cost-effectiveness studies (i.e. relevant articles as defined above), data extraction covering the following information will be 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
• Effectiveness
• Costs
• Incremental cost-effectiveness ratio (ICER)

4.4.1.5 Synthesis of findings
Like in Snowshill et al., the characteristics and main results of the included trial- and model-based studies will be described. The narrative synthesis provided by Snowsill et al. will be updated by including the evidence published from 2017 onwards.

4.4.2 Cost-effectiveness analysis
A Swiss cost-effectiveness analysis based on the MISCAN model was recently published.76 A de novo cost-effectiveness analysis would require unreasonable time and resources, therefore adapting the existing model is considered the most efficient approach. This MISCAN-based analysis was calibrated using individual level-data from the NLST and the Prostate, Lung, Colorectal, and Ovarian Cancer (PLCO) Screening Trial. The model was already adapted using Swiss specific input parameters regarding lung cancer epidemiology, smoking behaviours, and treatment costs. The effects and costs of 648 screening scenarios with different screening start and stop ages, smoking eligibility criteria, and screening intervals were examined from a public healthcare system perspective across a lifetime horizon in a cohort born between 1935 and 1965. The results suggested that several screening scenarios may be cost-effective in Switzerland, showing an ICER below CHF 50,000 per life-year gained.

The model will be updated with the effectiveness results of the NELSON trial (instead of NLST) and, if feasible, an update of all other Swiss input parameters (ranging from smoking behavior in Switzerland to lung cancer epidemiology and treatment costs). The improvement of the model will consist in implementing QALY estimates in the model and, if possible, potential effects of smoking cessation programs.

Since the use of the MISCAN model is covered by copyright, and since its structure cannot be freely changed, a close collaboration with the team of Prof. Harry de Koning of the Erasmus Medical Center in Rotterdam will be indispensable. The collaboration offers two major advantages. First, the MISCAN model is one of the most recognized lung cancer screening models developed on behalf of the U.S. National Cancer Institute. Second, Prof. Harry de Koning is the principal investigator of the NELSON trial, and hence, collaborating with the research group responsible for the NELSON trial may allow getting additional or unpublished trial information. The recently published 10 years results of the NELSON trial will contribute substantially to the update of the MISCAN model.

Prof. de Koning’s team has expressed interest in collaborating with us on this project.

4.4.3 Cost-benefit and budget impact analyses
The aim of the cost-benefit and budget impact analyses (BIA) will be to investigate the economic impact of different LDCT screening strategies in comparison to no screening (or the current situation) in Switzerland. The overall costs of LDCT screening (in particular of LDCT and follow-up tests) will be compared to potential economic benefits (e.g. in terms of reduced treatment costs or productivity loss).
4.4.4  Perspective
Costs will be assessed from a third-party payers perspective. A societal perspective will be added only if possible.

4.4.5  Subgroup and sensitivity analyses
The cost-effectiveness analyses as well as the BIA will provide estimations for several study populations according to the screening inclusion criteria. This will automatically lead to cost and cost-effectiveness estimations for different subgroups according to smoking intensity, start and stop age of screening, and number of years since smoking cessation. If technically feasible, potential differences between gender will be investigated.

In the sensitivity analyses, parameter uncertainty will be addressed by varying sets of related inputs. These parameters include attendance rates, discounting rates, LDCT costs, and treatment costs. Attendance rate in the base case will be assumed to be 100%, whereas in the sensitivity analyses low (35%), average (50%), and high (65%) attendance rates will be assumed. Discounting, originally set at 3%, will be varied from 0% to 6%. Finally, LDCT and treatment costs will be varied by 30% if compared to the base-case estimations.

In case of a new cost-effectiveness analysis including the potential impact of smoking cessation therapies, a variation of the therapy effectiveness will be considered.
5 Ethical considerations

5.1 Context
The prospect of LDCT screening for lung cancer raises several ethical issues that require careful consideration. The ethical analysis in this HTA will be based upon the conclusions regarding clinical effectiveness and cost-effectiveness, and will consider patient and public perspectives on screening. The ethical issues listed in Table 5 are formulated very general and will be refined based on the findings from the two previous sections.

5.2 Ethical, legal and social issues
The following ethical issues will be included in the analysis (Table 5). Most of these issues concern those who might be invited to screening, but some concern the general population in Switzerland.

Table 5: Ethical issues according to population or public perspectives

<table>
<thead>
<tr>
<th>Perspective of Population invited to screening</th>
<th>Following ethical issues will be addressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Ethical issues concerning informed consent and shared decision making, particularly discussing and communicating potential risk and prospective benefits of screening with those invited to be screened (including importance of absolute vs relative risk).</td>
<td></td>
</tr>
<tr>
<td>- Risk/benefit analysis from the perspective of the potentially screened person: potential clinical benefit vs false positives/distress/harm/side effects/incidental findings.</td>
<td></td>
</tr>
<tr>
<td>- Issues in recruitment to screening – ‘reluctant’ patient population, stigmatisation.</td>
<td></td>
</tr>
<tr>
<td>- Ethical issues in screening modalities: intervals, travel to hospital, radiation exposure, and smoking cessation.</td>
<td></td>
</tr>
<tr>
<td>- Further ethical issues may be added if deemed relevant during assessment.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wider public</th>
<th>- Issues in ethical resource allocation: does LDCT screening represent ethical/fair distribution of scarce resources? (Also covering the issue of who pays for screening).</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>- Public perspectives on justice – for example, should those suffering from what is perceived as ‘self-inflicted’ disease be allocated resources funded by the whole population? (particularly given cost-effectiveness questions).</td>
</tr>
<tr>
<td></td>
<td>- Further ethical issues may be added if deemed relevant during assessment.</td>
</tr>
</tbody>
</table>

5.3 Methods

5.3.1 Sources
The ethical analysis will be informed by:

- a systematic literature review by searching Medline via PubMed and additional purposive sampling of both ethics journals and medical journals, including qualitative research on public perspectives.
- the findings of the clinical effectiveness and cost-effectiveness assessment of this HTA report, which will also provide data on the specific Swiss context.
- and only if insufficient literature is found - focus groups or public workshops to generate additional data on patient and public perspectives.

5.3.2 Ethical analysis
Ethical analysis of each of the issues identified in Section 5.2 and any others emerging from the literature review and/or the conclusions of the previous parts of the HTA regarding clinical effectiveness and cost-effectiveness. Each issue will be subjected to thorough normative analysis, applying the main principles of biomedical ethics (respect for autonomy, beneficence, nonmaleficence and justice)\textsuperscript{110,111} and public health ethics.\textsuperscript{112}
6 References


