Screening for Lung Cancer:
Systematic Review and Meta-analyses

Final Submission:

March 31, 2015

McMaster Evidence Review and Synthesis Centre Team:

Leslea Peirson, Muhammad Usman Ali, Rachel Warren, Meghan Kenny

Maureen Rice, Donna Fitzpatrick-Lewis, Diana Sherifali, Parminder Raina

McMaster University, Hamilton Ontario Canada

Evidence Review Clinical Expert:

Dr. John Miller

Canadian Task Force on Preventive Health Care Working Group:

Gabriela Lewin (Chair), Maria Bacchus, Neil Bell, Jim Dickinson, Harminder Singh

Public Health Agency of Canada Scientific Research Managers:

Lesley Dunfield and Alejandra Jaramillo Garcia

Abstract

Background: This report was produced for the Canadian Task Force on Preventive Health Care (CTFPHC) to inform the development of guidelines on the screening of adults for lung cancer. The last CTFPHC guideline on this topic was published in 2003.

Purpose: To synthesize evidence on the benefits and harms of screening asymptomatic adults who are at average and high risk for lung cancer.

Data Sources: For benefits of screening we searched CENTRAL, Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R), and Embase from May 2012 to May 13, 2014 to update the search conducted for the Cochrane 2013 review on this same topic. The same databases were searched to look at harms of screening but the date range was extended to 2000. We also searched for evidence to answer the contextual questions (Embase and MEDLINE; 2009-June 2014), checked reference lists of included studies and relevant systematic reviews, and conducted a targeted grey literature search.

Study Selection: The titles and abstracts of papers considered for the key questions and sub-questions were reviewed in duplicate; any article marked for inclusion by either team member went on to full text screening. Full text screening was done independently by two people with consensus required for inclusion or exclusion. For benefits we included randomized controlled trials (RCTs) of screening interventions using chest x-ray (CXR), sputum cytology (SC) and/or low-dose computed tomography (LDCT) in adult populations that reported lung cancer mortality, all-cause mortality, smoking cessation rates, stage at diagnosis or incidental findings. All studies reporting harms of screening or invasive follow-up testing (i.e., overdiagnosis; death, major complications or morbidity from invasive follow-up testing; false positives and consequences of false positives; negative consequences of incidental findings; anxiety; quality of life; infection or bleeding from invasive follow-up testing) were included, regardless of design.

Data Abstraction: Review team members extracted data about the population, study design, intervention, analysis and results for outcomes of interest. One team member completed full abstraction, followed by a second team member who verified all extracted data and ratings. We assessed study quality using Cochrane's Risk of Bias tool (RCTs) and the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) framework. For the contextual questions, inclusion screening and abstraction were done by one person.

Analysis: Risk ratios (RRs) and 95% confidence intervals (CIs) for binary outcomes of benefits of lung cancer screening were calculated using random-effects models. Binary outcomes of harms of screening were reported using proportion per 1,000. Continuous outcomes of harms of screening (e.g., anxiety and quality of life) were reported as mean difference or mean change scores with 95% CIs. Test properties were reported descriptively using means or medians with ranges. GRADE tables were prepared for critical benefits (lung cancer and all-cause mortality) and critical harms (overdiagnosis, death or major complications/morbidity resulting from

invasive follow-up testing). For all other outcomes and subgroups, available data were meta-analyzed when appropriate or presented narratively.

Results: Thirty-three studies formed the evidence base for this review; 13 RCTs were used to answer the question regarding the benefits of screening for lung cancer and 30 studies provided data to answer the question about harms of screening or invasive follow-up testing. For the critical outcomes of lung cancer mortality and all-cause mortality, the low GRADE quality evidence indicated there is no benefit of CXR screening, with or without SC, when compared to no screening or less intensive screening. Pooled analyses of preliminary results from three relatively small trials comparing LDCT to usual care in high risk adults found no significant benefits for mortality with five years or less follow-up. One high quality trial with a large sample of high risk adults (NLST) and a median follow-up of 6.5 years found screening with LDCT showed significant benefits for mortality when compared with screening with CXR [lung cancer mortality RR 0.80 (95% CI 0.70, 0.92), NNS 308 (95% CI 201, 787); all-cause mortality RR 0.94 (95% CI 0.88, 1.00), NNS 219 (95% CI 115, 5,556)]. Two studies that examined subgroups of interest (age, gender, smoking history) found no differences in lung cancer mortality between the CXR screened and unscreened participants. For the important outcome of stage at diagnosis, most screening strategies for lung cancer showed statistically significant benefits in terms of disease detection. For CXR and LDCT screening, more cases of early stage non-small cell lung cancer and fewer cases of late stage malignancy were observed in the screened and more intensively screened groups compared to the control groups; with the exception of early stage disease detection using dual testing with CXR and SC compared to less intensive screening. In one large trial (NSLT), LDCT demonstrated better efficacy than CXR, detecting significantly more cases of early stage disease (57.0% versus 39.1%) and significantly fewer cases of late stage disease (43.0% versus 60.9%). Though a limited pool of evidence was available, none of the studies that reported on smoking cessation rates found a difference between screened and control groups. Lung cancer screening tests detect a variety of other clinically significant abnormalities, however little and inconsistent evidence was found regarding incidental findings of lung cancer screening.

The evidence for harms was primarily obtained from observational studies resulting in low GRADE quality of evidence. CXR screening was associated with: overdiagnosis ranging from 2.27% to 16.28%; 28.60 deaths (95% CI 16.02, 41.17) and 63.32 patients with major complications (95% CI 42.92, 92.49) per 1,000 patients undergoing invasive follow-up testing; median false positives of 65.0 (range 34.0 to 136.7) per 1,000 adults screened; and 2.30 (95% CI 1.49, 3.11) and 2.73 (95% CI 0.96, 4.51) individuals per 1,000 screened, who had benign conditions were subjected to minor and major invasive procedures, respectively, as part of diagnostic follow-up. LDCT screening was associated with: overdiagnosis ranging from 10.99% to 25.83%; 11.18 deaths (95% CI 5.07, 17.28) and 43.29 patients with major complications (95% CI 32.00, 54.58) per 1,000 patients undergoing invasive follow-up testing; median false positives of 167.1 (range 79.0 to 255.3) per 1,000 participants tested with a baseline or a single LDCT screen and 233.0 (95% CI 6.4, 690.0) per 1,000 screened with multiple rounds of testing; and 7.16 (95% CI

3.27, 11.05) and 4.98 (95% CI 3.68, 6.29) individuals per 1,000 screened, who had benign conditions were subjected to minor and major invasive procedures, respectively, as part of diagnostic follow-up. Little and inconsistent evidence was available regarding the other critical and important harms of interest (anxiety, quality of life, infection and bleeding from invasive follow-up testing).

Recent evidence was located to address a number of contextual questions. LDCT test properties varied across studies depending on the type of reference standard applied, cut-off or threshold value for a positive test, and LDCT technology and technique used. Sensitivity ranged from 80% to 100% and specificity ranged from 28% to 100%. No evidence was found that directly compared LDCT technologies or protocols used by radiologists and their effect on test performance. Diagnostic test properties were highest overall with a multi slice detector, computer assisted reading/diagnosis, and two independent radiologist readers. High risk adults indicated high willingness to be screened for lung cancer, reported neutral or positive screening experiences, and identified some individual and health care system barriers to screening. Variations in burden of lung cancer among Canadian rural, remote, Aboriginal and other ethnic populations are largely a reflection of tobacco use rates among those groups. There is no recent evidence on differential performance of lung cancer screening tests by these subpopulations. Modelling studies suggest annual LDCT screening is the most effective strategy, increasing diagnoses of lung cancer at earlier and more treatable stages and reducing rates of overdiagnosis. The absolute costs of lung cancer screening are difficult to estimate because of diversity across health care systems, variety in outcomes, and different assumptions about hypothetical cohorts.

Limitations: Although meeting our inclusion criteria, there was substantial variability across studies in terms of sample characteristics, screening tests, outcomes, comparators, length of follow-up, locations and timing. Only a few studies reported on some of the important outcomes and no evidence was found for a couple important harms. Only two papers included analyses to address the question about sub-group differences. Most of the harms data was obtained from observational studies. Publication bias could not be evaluated, given the low number of included studies in meta-analyses. Test properties were only examined for LDCT screening. No lung cancer risk assessment tools were located. Only papers in English or French were considered.

Conclusion: Considering mortality outcomes the available evidence indicated there is no benefit of CXR screening, with or without SC, when compared to no screening or less intensive screening in average to high risk adults. When compared with CXR in high risk individuals, LDCT reduced lung cancer mortality by 20% and all-cause mortality by 6%. A number of on-going LDCT trials will provide more conclusive evidence on the effectiveness of LDCT screening and further insights into optimal age for screening, screening interval and frequency. It is important to acknowledge that the poor specificity and the associated harms of LDCT pose challenges for clinicians and public health professionals to implement screening and warrant the need to develop standardized practices.

PROSPERO Registration #: CRD42014009984

Table of Contents

List of Acronyms

ARR	Absolute risk reduction
CI	Confidence interval
COPD	Chronic obstructive pulmonary disease
CQ	Contextual question
CTFPHC	Canadian Task Force on Preventive Health Care
CXR	Chest radiography (x-ray)
DANTE	Detection and Screening of Early Lung Cancer by Novel Imaging Technology and Molecular Essays Trial
DLCST	Danish Lung Cancer Screening Trial
EQ-5D	EuroQOL five dimensions questionnaire
GRADE	Grading of Recommendations, Assessment, Development and Evaluations
HR	Hazard ratio
IDEAL	Iressa Dose Evaluation in Advanced Lung Cancer
LDCT	Low-dose computed tomography
KQ	Key question
MCS	Mean change score
MD	Mean difference
MILD	Multi-centric Italian Lung Detection Trial
NELSON	Nederlands-Leuvens Longkanker Screenings Onderzoek Study
NLST	National Lung Screening Trial
NNS	Number needed to screen
NPV	Negative predictive value
NSCLC	Non-small cell lung cancer
OR	Odds ratio
PLCO	Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial
PPV	Positive predictive value
PROSPERO	International Prospective Registry of Systematic Reviews
QALY	Quality-adjusted life-year
RCT	Randomized controlled trial
RR	Risk ratio
SC	Sputum cytology
SES	Socio-economic status
SF-12	12-item Short-Form questionnaire
TVDT	Tumor volume doubling time
UK	United Kingdom
UKLS	United Kingdom Lung Screening Study
US	United States
USPSTF	United States Preventive Services Task Force

Chapter 1: Introduction

Purpose and Background

This report will be used by the Canadian Task Force on Preventive Health Care (CTFPHC) to inform a 2015 update of its 2003 guidelines on screening adults for lung cancer.[1] This systematic review synthesizes the benefits and harms of lung cancer screening in average and high risk asymptomatic adults and answers a number of contextual questions that consider issues including test properties and performance and participants' preferences regarding screening for lung cancer.

Definition

Lung cancer is a form of cell malignancy that begins in the lungs. Non-small cell lung cancers [(NSCLC) e.g., adenocarcinoma, squamous cell carcinoma, and large cell carcinoma] are the most common sub-types of the disease; more rarely diagnosed are the faster-growing small cell lung cancers (e.g., small cell carcinoma, mixed small cell/large cell, and combined small cell carcinoma).[2,3] This review deals primarily with the NSCLCs.

Prevalence and Burden of Lung Cancer

Lung cancer is estimated to be the most commonly diagnosed form of cancer in Canada (estimated 25,500 new cases in 2013) as well as the main cause of cancer related mortality among Canadians (estimated 20,200 deaths attributed to lung cancer in 2013).[4] Almost all (97%) of the estimated new cases of lung cancer in 2013 are expected to be identified in adults aged 50 years and older.[4] For the same year, the age-standardized incidence of lung cancer in men is estimated at 60.1 cases per 100,000 compared with 46.8 cases per 100,000 in women.[4] While the incidence is currently higher in men than women, the incidence for men became stable about 30 years ago (approximately 20 years after a reduction in smoking prevalence among men) and has significantly ($P<0.01$) decreased each year since the late 1990s; whereas the incidence for women has been increasing steadily ($P<0.01$) and has not yet reached a similar plateau following a general decline in tobacco consumption in the mid-1980s.[5] Lung cancer has a poor prognosis and the five-year relative survival rate is among the lowest for all types of cancer in Canada (17% in 2013).[4]

Risk Factors

Cigarette smoking is the main risk factor for developing lung cancer, and is associated with over 85% of the cases of this disease in Canada.[6] The *Canadian Tobacco Use Monitoring Survey* reported 44% of adults (4.6 million Canadians) were current or ever smokers (16% were current smokers) in 2012.[7] Other factors that increase risk for lung cancer include second hand exposure to cigarette smoke, exposure to radon and other toxic substances (e.g., asbestos, arsenic, diesel exhaust, silica, and chromium), having a first degree relative with lung cancer, and undergoing radiation therapy to the chest.[6,8]

Previous CTFPHC Recommendations and Recommendations from Other Guideline Developers

Updating its 1994 guidelines on lung cancer screening, in 2003 the CTFPHC determined that there was fair evidence upon which to recommend against using CXR to screen asymptomatic individuals for lung cancer, and insufficient evidence to inform a recommendation for or against using LDCT as a screening test for asymptomatic adults.[1] In 2004 the United States Preventive Services Task Force (USPSTF) concluded there was insufficient evidence to recommend for or against screening asymptomatic persons for lung cancer using either CXR or LDCT.[9] Newly published mortality results from the NLST appear to have convinced guideline groups across North America to rethink their recommendations regarding lung cancer screening.[10] The USPSTF's recently (2013) updated recommendation now endorses annual screening using LDCT for older adults (aged 55 to 80 years) who are current or former (quit within last 15 years) smokers with a minimum 30 pack-year smoking history (one pack=20 cigarettes; pack-year=daily consumption of one pack per day for one year; smoking two packs per day for one year would count as two pack-years).[11] Lung cancer screening using LDCT for similar high risk groups is also currently recommended by several other US organizations including the American Cancer Society,[12] the American College of Chest Physicians,[13] the American Lung Association,[14] the American Association for Thoracic Surgery,[15] and the National Comprehensive Cancer Network.[16] Likewise, in 2013 Cancer Care Ontario issued new guidelines recommending the use of LDCT to screen asymptomatic high risk adults for lung cancer.[17]

Scan of New Evidence since Previous Recommendation

Results from the NLST trial were first published in 2011.[10] This trial compared screening with LDCT to screening with CXR in a sample of high risk adults and showed a 20% relative reduction in lung cancer mortality for LDCT over a median follow-up of 6.5 years.[10] Several other lung cancer screening trials are underway and have published preliminary results for LDCT testing, although they have not shown the same mortality benefit as observed by the NLST.[18-24] Systematic reviews on the benefits of screening for lung cancer using LDCT have been published, including a 2013 Cochrane review[25] and the systematic review that supported the most recent USPSTF recommendation.[26]

Chapter 2: Methods

Review Approach

This review incorporates studies included in the 2013 Cochrane[25] and USPSTF[26] reviews on this same topic, and updates the search for benefits of lung cancer screening conducted for the 2013 Cochrane review.[25] A new search was conducted for harms of lung cancer screening to ensure all literature reporting harms ranked as critical would be identified. The review was developed, conducted and prepared according to the CTFPHC methods (http://canadiantaskforce.ca/methods/methods-manual/). The protocol was registered with the International Prospective Registry of Systematic Reviews (PROSPERO #CRD42014009984).

Analytic Framework, Key Questions and Contextual Questions

The analytic framework for this review is presented in Figure 1.

The key questions (KQ) and sub-questions considered for this review are:

KQ1. What are the clinical benefits of screening for lung cancer in adults not suspected of having lung cancer (lung cancer mortality, all-cause mortality, stage at diagnosis, smoking cessation rate, incidental findings)?
    a. What is the difference in screening effectiveness in populations and subgroups with varying risk for lung cancer (age, gender, smoking history)?

KQ2. What are the harms of screening for lung cancer in adults not suspected of having lung cancer (overdiagnosis, death from invasive follow-up testing, major complications or morbidity from invasive follow-up testing, false positives, consequences of false positives, negative consequences of incidental findings, anxiety, quality of life, infection from invasive follow-up testing, bleeding from invasive follow-up testing)?
    b. What is the difference in harms in populations and subgroups with varying risk for lung cancer (age, gender, smoking history)?

The contextual questions (CQ) considered for this review are:

CQ1. What is the evidence that test characteristics for effective lung cancer screening tests (sensitivity and specificity, false positives and false negatives, negative and positive predictive values, and test positivity rate) differ by subgroups with varying risk for lung cancer?

CQ2. What is the difference in test performance with changes and improvements in low-dose computed tomography technology or varying protocols used by radiologists?

CQ3. What are participants' values and preferences on screening for lung cancer?

CQ4. What is the optimal screening interval for screening for lung cancer?

CQ5. What risk assessment tools are identified in the literature to assess the risk of lung cancer?

CQ6. What is the evidence that subgroups (Aboriginal populations, rural or remote populations, other ethnic populations) have a higher burden of disease, a differential treatment response, differential performance of screening tests, or barriers to implementation?

CQ7. What is the cost-effectiveness of screening for lung cancer?

Search Strategy

For the key question on benefits of screening for lung cancer we updated the search conducted for the 2013 Cochrane review on lung cancer screening (same databases and search terms).[25] We searched CENTRAL, Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R), and Embase from May 2012 (the date of the last Cochrane search) to May 13, 2014 for RCTs on screening for lung cancer published in English or French. The same databases were searched to look at harms of screening but the date range was extended to 2000 and no limits were placed on study design. Reference lists of on topic systematic reviews were searched for relevant studies not captured by our search. A separate search was conducted to look for evidence that would answer the contextual questions; this strategy included two databases (Embase and MEDLINE) and covered the period between January 2009 and June 12, 2014. A focused web-based grey literature search of Canadian sources was also undertaken for the contextual questions. The full search strategies are provided in Appendix 1.

Study Selection

After removing duplicates, all citations found through our updated search were uploaded to a web-based systematic review software program[27] for screening. In addition, the studies included in the 2013 lung cancer screening reviews by the Cochrane group[25] and the USPSTF,[26] as well as the studies identified through hand-searching systematic review reference lists, were added to the pool of citations available for relevance testing against the inclusion criteria for this review.

Titles and abstracts of papers considered for the key questions and sub questions were reviewed in duplicate; articles marked for inclusion by either team member went on to full text relevance testing. Full text screening was done independently by two people with consensus required for inclusion or exclusion.

For citations located in the contextual questions search, title and abstract screening was done by one person.

Inclusion and Exclusion Criteria

Language

The published results of studies had to be available in either English or French.

Population

The population of interest for this review is asymptomatic adults aged 18 years and older who are at average or high risk but are not suspected of having lung cancer (e.g., may have a cough). The population includes current, former and second-hand smokers, as well as those with exposures to substances that may affect risk and other identified factors that may increase risk.

Excluded from this review are studies that focused on people under age 18 or that targeted adults 18 years and older who were either suspected of having lung cancer or were previously diagnosed with lung cancer.

Interventions

The three lung cancer screening interventions of interest included: (1) chest radiography (CXR), (2) low-dose computed tomography (LDCT), and (3) sputum cytology (SC).

Study Design and Comparison Groups

To answer the key question about the benefits of screening, only RCTs with comparison groups of no screening or comparison between tests were eligible for inclusion; case control, case series and ecological studies were excluded. Any quantitative study design (with or without comparison groups) was considered acceptable to answer the key question about harms of screening.

Outcomes

A CTFPHC working group identified and ranked benefits and harms of lung cancer screening as critical, important or not important in terms of their importance for guideline decision making, using the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) approach.[28] GRADE ratings from 7 to 9 indicate critical outcomes and ratings from 4 to 6 are considered important. Outcomes with ratings from 1 to 3 are not considered important, and thus are not examined in this review. For the key question about the benefits of lung cancer screening, the two critical outcomes of interest are lung cancer mortality (GRADE rating=9) and all-cause mortality (9) and the three important outcomes are smoking cessation rate (6), stage at diagnosis (6) and incidental findings (e.g., diagnosis of a thoracic aneurysm) (6). For the key question about the harms of lung cancer screening or follow-up testing, the three critical outcomes are overdiagnosis (9), death from invasive follow-up testing (9), and major complications or morbidity as a result of invasive follow-up testing (7), and the seven important outcomes are false positives (6), consequences of false positives (6), negative consequences of incidental findings (6), anxiety (5), quality of life (5), infection from invasive follow-up testing (5), and bleeding from invasive follow-up testing (5).

Data Abstraction

For each study used to answer the key questions, review team members extracted data about the population, the study design, the intervention, the analysis and the results for outcomes of interest. To answer the key question about benefits we extracted the number of events data for the longest available follow-up if multiple follow-up points were provided. To answer the key question about harms we extracted data for reported adverse events of interest, only if they were attributed to lung cancer screening or invasive follow-up testing. In addition, for the analyses we only included mutually exclusive adverse events data, that is, we selected results that reported the number of participants who experienced at least one event in the respective overall adverse effects category. The results from studies that reported the total number of adverse events experienced across all study group participants are captured only in the narrative results of this review. For each study, one team member completed full abstraction (study characteristics, risk of bias assessment, outcome data) using electronic forms housed in a web-based systematic review software program.[27] A second team member then verified all extracted data and ratings;

disagreements were resolved through discussion and/or third party consultation when consensus could not be reached.

Assessing Risk of Bias

Arriving at a GRADE rating for a body of evidence (see next section) requires a preliminary assessment of the risk of bias or study limitations for the individual studies.

All RCTs included to answer the benefits of screening question were assessed using the Cochrane Risk of Bias tool.[29] This rating tool covers a number of domains: sequence generation; allocation concealment; blinding; incomplete outcome reporting; selective outcome reporting; and other risk of bias. We separated our assessment of blinding for objectively assessed (lung cancer mortality, all-cause mortality, stage at diagnosis, incidental findings) and self-report (smoking cessation) outcomes. For other sources of risk of bias we selected: industry funding (without an explicit statement that sponsors were not involved in other aspects of the study or resultant publications), low sample size (<30 participants per arm), contamination of the control group through opportunistic screening (≥20% receiving screening tests), and significant baseline differences between study groups on factors that might affect outcomes of interest.

Information to determine risk of bias was abstracted from the primary methodology paper for each study and any other relevant published papers. For each study, one team member completed the initial ratings which were then verified by a second person; disagreements were resolved through discussion and/or third party consultation when consensus could not be reached. To assign a high or low risk of bias rating for a particular domain we looked for explicit statements or other clear indications that the relevant methodological procedures were or were not followed. In the absence of such details we assigned unclear ratings to the applicable risk of bias domains. To determine the overall risk of bias rating for an outcome we considered all domains, however greater emphasis was placed on assessments of the first three areas of randomization, allocation and blinding because these represent the most significant sources of introducing bias to a randomized controlled trial and hence could lead to biased estimates and conclusions.[29]

Table 1 summarizes the risk of bias ratings applied to the RCTs included to answer the key question about the benefits of screening.

The observational studies used to answer the key question regarding harms were not systematically assessed for methodological quality.

There was no assessment of the methodological quality of studies used to answer the contextual questions (http://canadiantaskforce.ca/methods/methods-manual/).

Assessing Strength or Quality of the Evidence

The strength of the evidence was determined based on the GRADE system of rating the quality of evidence using GRADEPro software.[30, 31] This system of assessing evidence is widely used and is endorsed by over 40 major organizations including the World Health Organization, Centers for Disease Control and Prevention, and the Agency for Healthcare Research and

Quality.[32] The GRADE system rates the quality of a body of evidence as high, moderate, low or very low; each of the four levels reflects a different assessment of the likelihood that further research will impact the estimate of effect (i.e., high quality=further research is unlikely to change confidence in the estimate of effect; moderate quality=further research is likely to have an important impact on confidence in the estimate of effect and may change the estimate; low quality=further research is very likely to have an important impact on confidence in the estimate of effect and is likely to change the estimate; very low quality=the estimate of effect is very uncertain).[32]

A GRADE quality rating is based on an assessment of five conditions: (1) risk of bias (limitations in study designs), (2) inconsistency (heterogeneity) in the direction and/or size of the estimates of effect, (3) indirectness of the body of evidence to the populations, interventions, comparators and/or outcomes of interest, (4) imprecision of results (few participants/events/observations, wide confidence intervals), and (5) indications of reporting or publication bias. Grouped studies begin with a high quality rating which may be downgraded if there are serious or very serious concerns across the evidence related to one or more of the five conditions.

Full GRADE assessments were conducted only for the two benefit outcomes that the CTFPHC working group rated as critical (lung cancer mortality and all-cause mortality). Data were entered into the GRADEPro software along with the quality assessment ratings to produce two analytic products for each critically ranked outcome and the comparisons of interest: (1) a GRADE Evidence Profile Table and (2) a GRADE Summary of Findings Table.

Data Analysis

To perform meta-analyses for the critical and important benefits of lung cancer screening (lung cancer mortality, all-cause mortality and stage distribution), we utilized the number of events, proportion or percentage data from included RCTs to generate the summary measures of effect in the form of risk ratios (RR) using the DerSimonian and Laird random effects model with inverse variance method.[33] When assessing the stage distribution it is important to evaluate both the proportion of early (stage I and II) and late stage (III and IV) cancers, therefore the total number of lung cancers detected in each study was taken as the denominator. The random effects model assumes the studies are a sample of all potential studies and incorporates an additional between-study component to the estimate of variability. The primary subgrouping in these meta-analyses was based on the type of screening test(s) used (CXR, CXR plus SC, LDCT). Cochran's Q ($\alpha=0.05$) and I^2 ($\geq75\%$=substantial heterogeneity) statistics were employed to quantify statistical heterogeneity between studies.

In addition, for the critical benefits of lung cancer screening (lung cancer mortality and all-cause mortality) that showed significant effects, we calculated absolute risk reduction (ARR) and number needed to screen (NNS) and added these values to the GRADE tables. NNSs were calculated using the absolute numbers computed by the GRADE software.[30] GRADE estimates

the absolute number per million using the control group event rate and risk ratio with the 95% confidence interval obtained from the meta-analysis.

For harms of lung cancer screening with binary data such as major complications or morbidity and death due to invasive follow-up testing, over-diagnosis, false positives, consequences of false-positives, and infections, the number of events, proportion or percentage data was utilized to generate the summary measures of effect using the DerSimonian and Laird random effects model with inverse variance method.[33] The binomial confidence intervals for each proportion/rate were calculated using the Wilson score interval method.[34]

For harms of lung cancer screening with continuous data such as anxiety and quality of life, the summary measures of effect were generated in the form of mean difference (MD) between intervention and control groups using the DerSimonian and Laird random effects model with inverse variance method.[33] If data for the control group were not available, the results were synthesized descriptively using change from baseline (i.e., difference between pre-screening and post-screening) or mean change score (MCS) for intervention group only and reported as MCS with 95% CI.

The analyses were performed using Review Manager version 5.3 software,[30, 35] STATA version 12[36] and GRADEpro.[30]

When data for particular outcomes (e.g., incidental findings) were inconsistently reported across studies or when studies did not provide data necessary for pooling (e.g., reported only a P-value or did not report values for the control group), the results are described narratively.

Chapter 3: Results

Summary of the Literature Search for Key Questions

Our updated search for studies examining the benefits of lung cancer screening and our extended search for studies reporting harms of screening, located 840 unique citations (Figure 2). From these searches we identified 29 on-topic systematic reviews, the reference lists of which were examined, resulting in 18 papers being added to our screening database. We added 16 studies reported in the USPSTF review[26] and five studies included in the Cochrane review.[25] Together these sources yielded 879 studies that required title and abstract screening. Full text relevance testing took place on 223 studies. The majority of studies (161/223) was excluded for not meeting at least one key inclusion criterion (list of excluded studies available on the CTFPHC website http://canadiantaskforce.ca/) and another 29 were set aside as systematic reviews.

At the end of the search and selection process, 33 studies, many with multiple publications, met the inclusion criteria for this review. Thirteen of the studies addressed the key question on the benefits of screening and 30 provided data that were used to answer the key question on the harms of screening or follow-up testing.

In terms of the 13 studies identified for key question one, none were new or novel. Eight of the studies were also included in the USPSTF review;[10, 21, 23, 37-41] one of which was unique to this source.[41] Our updated search found publications with more recent data for four of the USPSTF identified studies.[21, 23, 37, 40] Three studies included by the USPSTF overlapped with studies included by the Cochrane group.[10, 38, 39] Five of our included studies were only found through the Cochrane review.[42-46]

Summary of the Included Studies

A total of 13 RCTs were included to answer the key question of the benefits of screening for lung cancer.[10, 21, 23, 37-46] Seven trials[10, 23, 37, 38, 40, 41, 44] included mixed gender samples and six trials[21, 39, 42, 43, 45, 46] included only men. There were 10 studies[10, 21, 23, 37, 39-43, 45] that enrolled only current and former smokers, two studies[38, 44] that targeted smokers and non-smokers, and one study[46] that included non-smokers, former smokers and current smokers. The range of ages included in the studies (upon enrollment) was 35 to 74 years; less than half of the studies[23, 40, 42, 43, 45, 46] included some participants <50 years of age and over half[10, 21, 37, 39-41, 43, 45, 46] included some participants >70 years of age. Two trials[38, 44] compared CXR to usual care, five trials[39, 42, 43, 45, 46] compared more intensive CXR (with or without SC) screening to less intensive screening, four trials[21, 23, 37, 40] compared LDCT screening to no screening or usual care, and two trials[10, 41] compared LDCT to CXR. Ten studies[10, 23, 37-40, 42-45] had follow-up of ≥5 years while three studies[21, 41, 46] reported follow-up of <5 years. Over half of the studies[23, 37, 39-42, 46] were rated as having an unclear risk of bias, primarily due to the lack of information about or lack of procedures to ensure random sequence generation, allocation concealment and blinding of outcome assessment (Table 1). Of the remaining six studies, three[10, 38, 45] were assigned a low risk of bias rating and three[21, 43, 44] were designated as high risk. Recruitment was initiated after the year

2000 in six studies,[10, 21, 23, 37, 40, 41] while one study[38] commenced recruitment in the 1990s, four studies[39, 42, 43, 45] in the 1970s, and two studies[44, 46] in the 1960s. Seven trials[10, 38, 39, 41, 43-45] were conducted in the US and all other trials[21, 23, 37, 40, 42, 46] were conducted in European nations (Denmark, the Netherlands, Belgium, Italy and Czechoslovakia). A high level summary of the body of evidence included to answer the key question about the benefits of lung cancer screening is provided in Table 2. The characteristics of the 13 RCTs are reported individually and in more detail in Table 3.

A total of 30 studies were included to answer the key question of the harms of screening for lung cancer; some of these studies overlap with trials included for the question about the benefits of screening.[10, 21, 23, 24, 37-42, 45, 47-65] Many of the studies providing harms data included multiple publications; Table 4 identifies these companion papers.

Results for Key Questions

KQ1. *What are the clinical benefits (B) of screening for lung cancer in adults not suspected of having lung cancer (lung cancer mortality, all-cause mortality, smoking cessation rate, stage at diagnosis, incidental findings)?*

Key findings across critical (mortality) and selected important (stage at diagnosis) outcomes, with pooled estimates of effect, are provided in Table 5. Detailed results for each outcome are presented below.

B-1.0 Lung Cancer Mortality

Eleven studies were identified that provided data on lung cancer mortality. Seven RCTs used CXR as the primary screening test;[38, 39, 42-46] in four of these studies CXR was combined with SC.[39, 42, 43, 45] Three RCTs used LDCT as the primary screening test.[21, 23, 40] One study compared LDCT with CXR.[10] Our search and selection process did not locate any RCTs that investigated SC as the only screening test that met our inclusion criteria, and reported on the outcome of lung cancer mortality.

Evidence Set 1 provides the GRADE Evidence Profile Table (1.1), the GRADE Summary of Findings Table (1.2), and the forest plots (1.1 and 1.2) generated for the outcome of lung cancer mortality.

B-1.1 Chest X-Ray (CXR)

Seven RCTs that used CXR as the primary screening test reported results for lung cancer mortality;[38, 39, 42-46] these studies were separated into three groups for analysis.

B-1.1.1 CXR Screening versus Usual Care

Two RCTs examined screening with CXR alone versus usual care (no formalized screening) (Kaiser,[44] PLCO[38]). These two studies have a combined sample of 165,575 (82,583 CXR; 82,992 usual care). Both studies included mixed gender populations with about equal representation of men and women. The Kaiser study targeted middle-aged adults (35 to 54 years)

and at outset the sample was about equally divided above and below age 45. The PLCO study targeted older adults (55 to 74 years) and enrolled about 33% of participants in their 50s, 53% in their 60s, and 13% in their early 70s. Both studies recruited smokers (current and former) as well as never smokers; only 17% of the Kaiser study sample were identified as smokers (though study authors suggested this was likely an underestimate of actual smokers) while in the PLCO study 52% were identified as current or former smokers. In the Kaiser study screening participants were offered CXR annually for four years while the PLCO participants were offered an annual CXR for 16 years as part of a comprehensive health check-up. The control group in both studies received usual care (screening was not offered or advised, but patients could be screened if they and/or their health care provider initiated testing). The length of follow-up for lung cancer mortality was up to 16 years in the Kaiser study and up to 13 years (median 11.9, mean 11.2, interquartile range 10 to 13) in the PLCO trial. Both studies were conducted in the US. The Kaiser study was initiated in 1964 while the PLCO trial started in 1993. There was no significant difference in lung cancer mortality between the group screened with CXR alone and the usual care group [RR 0.99 (95% CI 0.92, 1.07) I^2=0%]. This body of evidence was downgraded for concerns regarding risk of bias and imprecision; a low quality GRADE rating was assigned.

B-1.1.2 More Intensive CXR Screening versus Less Intensive Screening

The second comparison included a single RCT that examined more intensive screening using CXR against less intensive screening using the same test (North London Study[46]). This study included 55,034 participants (29,723 more intensive screening; 25,311 less intensive screening). Only men were included. The study targeted adults aged 40 to 70 years; about one-quarter of participants who were enrolled were aged 40 to 44, another quarter were aged 45 to 49, a bit less than a quarter (22%) were aged 50 to 54, and the remaining men were mostly in their late 50s (16%) or 60s (10%). At enrollment most of the sample (69%) were smokers, 19% were former smokers, and 12% were never smokers. Men were enrolled only if they showed no signs of lung cancer on a preliminary screen. Participants randomized to the more intense screening arm were then offered CXR every six months for a period of three years (up to six additional screens) while those assigned to the less intense screening arm had the eligibility CXR and then a second CXR three years later. The length of follow-up for lung cancer mortality was at least five years. This study was conducted in the UK during the 1960s. There was no significant difference in lung cancer mortality between the group that underwent more frequent CXR screens and the group that received less intensive testing [RR 1.03 (95% CI 0.74, 1.42)] This body of evidence was downgraded for concerns regarding risk of bias and imprecision; a low quality GRADE rating was assigned.

B-1.1.3 More Intensive CXR plus SC Screening versus Less Intensive Screening

The third group included four RCTs that examined screening using CXR plus SC versus less intensive screening (CXR and SC but fewer screens and longer interval; CXR alone; advised to have annual CXR and SC) (Czech Study,[42] Johns Hopkins Study,[43] Mayo Lung Project,[39] Memorial Sloan-Kettering Study[45]). These four studies have a combined sample of 35,983

(17,983 more intensive screening; 18,000 less intensive screening). All four studies included only men. At enrollment, about one-quarter of the Mayo Lung Project sample was under age 50, another quarter was aged 50 to 54, two-fifths were aged 55 to 65, and the remaining 10% were aged 65 and older. In the Czech study about 38% of the men were enrolled in their 40s, about half (48%) were in their 50s, and 14% were in their early 60s. About one-third of the Memorial Sloan-Kettering sample was under age 50 when recruited, a bit more than half (54%) of the men were in their 50s, and the remaining quarter were mostly in their 60s, with a small percentage (3%) in their 70s. In the Johns Hopkins study just under one-third (31%) of the men were in their late 40s, about a quarter (27%) were in their early 50s and another quarter (25%) were in their late 50s, and the remaining enrollees were mostly in their 60s. All four studies recruited only heavy smokers (current or recently quit). In the Mayo Lung study 94% of the men smoked $\geq$1 pack/day, 97% had smoked for $\geq$20 years and 91% had $\geq$25 pack-years. Similarly, all men in the Memorial Sloan-Kettering study smoked $\geq$1 pack/day and 91% had $\geq$25 pack-years. Likewise all men in the Johns Hopkins study smoked $\geq$1 pack/day and 94% had $\geq$25 pack-years. Half of the men in the Czech study had a lifetime consumption of 150,000 to 250,000 cigarettes (approximately 20.55 to 34.25 pack-years) and the other half had smoked more. Two studies only enrolled participants who showed no signs of lung cancer on a preliminary screen. In the Mayo Lung Project participants were offered CXR and SC every four months for a period of six years (up to 18 screens); the less intense screening arm had the eligibility CXR and SC test and was subsequently advised to get screened at least yearly, but these men were not offered systematic re-screening. The Czech study offered one group the dual screen every six months for three years (up to six screens) followed by three years of annual CXR testing; the less intense screening group received the dual test prior to randomization and again three years later, after which annual CXR was offered for three more years. The Memorial Sloan-Kettering and Johns Hopkins studies offered half of the men annual CXR and SC every four months for five years (initial recruits may have had an additional two or three years of screening), and the other half were offered annual screening with CXR alone for the same amount of time. In terms of follow-up, lung cancer mortality was assessed at a median of 20.5 years in the Mayo Lung Project, for up to 15 years in the Czech study, and for up to nine years in the Memorial Sloan-Kettering and Johns Hopkins studies. All four studies were initiated in the 1970s. Three studies were conducted in the US and one in Czechoslovakia. <u>There was no significant difference in lung cancer mortality between the group that received more intensive screening with CXR plus SC and the group that received less intensive screening</u> [RR 1.01 (95% CI 0.87, 1.18) I^2=59%]. This body of evidence was downgraded for concerns regarding risk of bias and imprecision; a low quality GRADE rating was assigned.

B-1.2 Low-Dose Computed Tomography (LDCT)

Four RCTs that used LDCT as the primary screening test reported results for lung cancer mortality;[10, 21, 23, 40] these studies were separated into three groups for analysis.

B-1.2.1 Annual LDCT versus Usual Care

Three RCTs examined annual screening with LDCT versus usual care (DANTE,[21] DLCST,[40] MILD[23]). These three studies have a combined sample of 9,489 (4,518 annual LDCT; 4,971 usual care). Two studies included mixed gender samples; slightly more than half of the participants in the DLCST study and about two-thirds of the MILD study sample were men. The DANTE study included only men. At enrollment, in the two mixed gender trials participants were in their late 50s (MILD median age 57 years; DLCST mean age 58 years); enrollees were slightly older (mean 64 years) in the male-only DANTE trial. All three studies recruited only current or former (quit in last 10 years) smokers. In the DANTE study a little over half of the men were current smokers and all the men had ≥20 pack-years (mean 47 pack-years). All participants in the DLCST study also had ≥20 pack-years (mean 36 pack-years) and the mean number of cigarettes smoked per day was 19. Smoking history was a bit heavier in the MILD study with 90% of control and 69% of screening participants smoking upon enrollment, a mean of 38 pack-years across participants with consumption for two-thirds of the group at ≥20 cigarettes per day. In one intervention arm of the MILD study, participants were offered annual LDCT screening (median of five screens); the control group received usual care. Likewise, participants in the intervention arm of the DLCST trial were offered five annual LDCT screens and the control group received usual care. At baseline, men in the screening group of the DANTE trial had a CXR plus SC test in addition to their first of five annual LDCT scans; the usual care group also had the baseline CXR and SC test (no LDCT) but then only attended yearly for clinical reviews (no screening). In terms of follow-up, lung cancer mortality was assessed at a median of 4.4 years (maximum six years) in the MILD study, for five years in the DLCST study and for a median of 2.8 years (range 1.8 to 79.2 months) in the on-going DANTE trial. All three studies were initiated in the last 10 to 15 years (since 2000). Two studies were conducted in Italy and one in Denmark. <u>There was no significant difference in lung cancer mortality between the group screened annually with LDCT and the usual care group</u> [RR 1.35 (95% CI 0.79, 2.29) I^2=32%]. This body of evidence was downgraded for concerns regarding risk of bias and imprecision; a low quality GRADE rating was assigned.

B-1.2.2 Biennial LDCT Screening versus Usual Care

One of the above mentioned RCTs included a second intervention arm to examine the effect of biennial (every two years) screening with LDCT versus usual care (MILD[23]). The MILD study included 2,909 participants (1,186 biennial LDCT; 1,723 usual care). The trial included mixed gender participants; about two-thirds of the sample was men. At enrollment, participants were mostly in their late 50s (median age 57 years). Only current or former (quit in last 10 years) smokers were recruited; 90% of control and 69% of screening participants were current smokers, the overall mean pack-years was 38 and consumption for two-thirds of the sample was ≥20 cigarettes per day. In the intervention arm participants were offered LDCT screening every two years (median of three screens); the control group received usual care. In terms of follow-up, lung cancer mortality was assessed at a median of 4.4 years (maximum six years). This trial was

initiated in 2005 in Italy. There was no significant difference in lung cancer mortality between the group screened biennially with LDCT and the usual care group [RR 1.25 (95% CI 0.42, 3.70)]. This body of evidence was downgraded for concerns regarding risk of bias and imprecision; a low quality GRADE rating was assigned.

B-1.2.3 LDCT Screening versus CXR Screening

A fourth RCT compared screening with LDCT with screening with CXR (NLST[10]). The NLST study included 53,454 participants (26,722 LDCT; 26,732 CXR). The trial included mixed gender participants; about 60% of the sample was men. At enrollment, about 60% of participants were over age 60. Only current (48%) or former (quit in last 15 years) (52%) smokers were included. Participants in each group were offered three annual screens with their respectively assigned test. In terms of follow-up, lung cancer mortality was assessed at a median of 6.5 years (maximum 7.4 years). This trial was initiated in 2002 in the US. The LDCT screening group showed a relative reduction of 20% in lung cancer mortality as compared to the CXR screening group [RR 0.80 (95% CI 0.70, 0.92); absolute value per million 3,250 fewer, range from 1,271 fewer to 4,972 fewer]. The absolute risk reduction (ARR) is 0.33%. The number needed to screen (NNS) to avert one death from lung cancer is 308 (95% CI 201, 787). This body of evidence was not downgraded on any domain; a high quality GRADE rating was assigned.

B-2.0 All-Cause Mortality

Nine studies were identified that provided data on all-cause mortality. Five RCTs used CXR as the primary screening test;[38, 39, 42, 44, 45] in three of these studies CXR was combined with SC.[39, 42, 45] Three RCTs used LDCT as the primary screening test.[21, 23, 40] One study compared LDCT with CXR.[10] Our search and selection process did not locate any RCTs that investigated SC alone as the screening test that met our inclusion criteria, and reported on the outcome of all-cause mortality.

Evidence Set 2 provides the GRADE Evidence Profile Table (2.1), the GRADE Summary of Findings Table (2.2), and the forest plots (2.1 and 2.2) generated for the outcome of all-cause mortality.

B-2.1 Chest X-Ray (CXR)

Five RCTs that used CXR as the primary screening test reported results for all-cause mortality;[38, 39, 42, 44, 45] these studies were separated into two groups for analysis.

B-2.1.1 CXR Screening versus Usual Care

Two RCTs examined screening with CXR alone versus usual care (no screening) (Kaiser,[44] PLCO[38]). These two studies have a combined sample of 165,575 (82,583 CXR; 82,992 usual care). Both studies included mixed gender populations with about equal representation of men and women. The Kaiser study targeted middle-aged adults (35 to 54 years) and at outset the sample was about equally divided above and below age 45. The PLCO study targeted older adults (55 to 74 years) and enrolled about 33% of participants in their 50s, 53% in their 60s, and 13% in their early 70s. Both studies recruited smokers (current and former) as well as never

smokers; only 17% of the Kaiser study sample were identified as smokers (though study authors suggested this was likely an underestimate of actual smokers) while in the PLCO study 52% were identified as current or former smokers. In the Kaiser study screening participants were offered CXR annually for four years while the PLCO participants were offered an annual CXR for 16 years as part of a comprehensive health check-up. The control group in both studies received usual care (screening was not offered or advised, but patients could be screened if they and/or their health care provider initiated testing). The length of follow-up for mortality outcomes was up to 16 years in the Kaiser study and up to 13 years (median 11.9, mean 11.2, interquartile range 10 to 13) in the PLCO trial. Both studies were conducted in the US. The Kaiser study was initiated in 1964 while the PLCO trial started in 1993. There was no difference in all-cause mortality between the group screened with CXR alone and the usual care group [RR 0.98 (95% CI 0.96, 1.00) I^2=0%]. This body of evidence was downgraded for concerns regarding risk of bias and imprecision; a low quality GRADE rating was assigned.

B-2.1.2 More Intensive CXR plus SC Screening versus Less Intensive Screening

The second group of three RCTs compared a dual screening test using CXR plus SC with less intensive screening (CXR and SC but fewer screens and longer interval; CXR alone; advised to have annual CXR and SC) (Czech study,[42] Mayo Lung Project,[39] Memorial Sloan-Kettering Study[45]). These three studies have a combined sample of 25,596 (12,757 more intense screening; 12,839 less intense screening). All three studies included only men. At enrollment, about one-quarter of the Mayo Lung Project sample was under age 50, another quarter was aged 50 to 54, two-fifths were aged 55 to 65 and the remaining 10% were aged 65 and older. In the Czech study about 38% of the men were enrolled in their 40s, about half (48%) were in their 50s, and 14% were in their early 60s. About one-third of the Memorial Sloan-Kettering sample was under age 50 when recruited, a bit more than half (54%) of the men were in their 50s, and the remaining quarter were mostly in their 60s, with a small percentage (3%) in their 70s. All three studies recruited only heavy smokers (current or recently quit). In the Mayo Lung study 94% of the men smoked at least one pack/day, 97% had smoked for ≥20 years and 91% had ≥25 pack-years. Similarly, all men in the Memorial Sloan-Kettering study smoked at least one pack/day and 91% had ≥25 pack-years. Half of the men in the Czech study had a lifetime consumption of 150,000 to 250,000 cigarettes (approximately 20.55 to 34.25 pack-years) and the other half had smoked more. Two studies only enrolled participants who showed no signs of lung cancer on a preliminary screen. In the Mayo Lung Project participants were offered CXR and SC every four months for a period of six years (up to 18 screens); the less intense screening arm had the eligibility CXR and SC test and was subsequently advised to get screened at least once a year, but these men were not offered systematic re-screening. The Czech study offered the experimental group the dual screen every six months for three years (up to six screens) followed by three years of annual CXR testing; the group receiving less intense screening received the dual test prior to randomization and again three years later, after which annual CXR was offered for three more years. The Memorial Sloan-Kettering study offered half of the men annual CXR and SC every four months for five to eight years, and the other half were offered annual

screening with CXR alone for the same amount of time. In terms of follow-up, all-cause mortality was assessed at a median of 20.5 years in the Mayo Lung Project, for up to nine years in the Memorial Sloan-Kettering study, and for six years in the Czech study. All three studies were initiated in the 1970s. Two studies were conducted in the US and one in Czechoslovakia. <u>There was no significant difference in all-cause mortality between the group screened with CXR plus SC and the group that received less intensive screening</u> [RR 1.04 (95% CI 0.97, 1.11) I^2=37%]. The GRADE rating for this body of evidence was low quality; downgrading occurred because of concerns regarding risk of bias and imprecision.

B-2.2 Low-Dose Computed Tomography (LDCT)

Four RCTs that used LDCT as the primary screening test reported results for all-cause mortality;[10, 21, 23, 40] these studies were separated into three groups for analysis.

B-2.2.1 Annual LDCT Screening versus Usual Care

Three RCTs examined annual screening with LDCT versus usual care (DANTE,[21] DLCST,[40] MILD[23]). These three studies have a combined sample of 9,489 (4,518 annual LDCT; 4,971 usual care). Two studies included mixed gender samples; slightly more than half of the participants in the DLCST study and about two-thirds of the MILD study sample were men. The DANTE study included only men. At enrollment, in the two mixed gender trials participants were in their late 50s (MILD median age 57 years; DLCST mean age 58 years); enrollees were slightly older (mean 64 years) in the male-only DANTE trial. All three studies recruited only current or former (quit in last 10 years) smokers. In the DANTE study a little over half of the men were current smokers and all the men had ≥20 pack-years (mean 47 pack-years). All participants in the DLCST study also had ≥20 pack-years (mean 36 pack-years) and the mean number of cigarettes smoked per day was 19. Smoking history was a bit heavier in the MILD study with 90% of control and 69% of screening participants smoking upon enrollment, a mean of 38 pack-years across participants with consumption for two-thirds of the group at ≥20 cigarettes per day. All three studies used annual LDCT as the screening test. In one intervention arm of the MILD study, participants were offered annual LDCT screening (median of five screens); the control group received usual care. Similarly, participants in the intervention arm of the DLCST trial were offered five annual LDCT screens; the control group received usual care. At baseline, men in the screening group of the DANTE trial had a CXR plus SC test in addition to their first of five annual LDCT scans; the usual care group also had the baseline CXR and SC test (no LDCT) but then only attended yearly for clinical reviews (no screening). In terms of follow-up, all-cause mortality was assessed at a median of 4.4 years (maximum six years) in the MILD study, for five years in the DLCST study and for a median of 2.8 years (range 1.8 to 79.2 months) in the on-going DANTE trial. All three studies were initiated in the last 10 to 15 years (since 2000). Two studies were conducted in Italy and one in Denmark. <u>There was no significant difference in all-cause mortality between the group screened annually with LDCT and the usual care group</u> [RR 1.42 (95% CI 0.91, 2.22) I^2=67%]. This body of evidence was downgraded for concerns regarding risk of bias and imprecision; a low quality GRADE rating was assigned.

B-2.2.2 Biennial LDCT Screening versus Usual Care

One of the above mentioned RCTs included a second intervention arm to examine the effect of biennial (every two years) screening with LDCT versus usual care (MILD[23]). This study included 2,909 participants (1,186 biennial LDCT; 1,723 usual care). The MILD trial included mixed gender participants; about two-thirds of the sample was men. At enrollment, participants were mostly in their late 50s (median age 57 years). Only current or former (quit in last 10 years) smokers were recruited; 90% of control and 69% of screening participants were current smokers, the overall mean pack-years was 38 and consumption for two-thirds of the sample was ≥20 cigarettes per day. In the intervention arm participants were offered LDCT screening every two years (median of three screens); the control group received usual care. In terms of follow-up, all-cause mortality was assessed at a median of 4.4 years (maximum six years). This trial was initiated in 2005 in Italy. There was no significant difference in all-cause mortality between the group screened biennially with LDCT and the usual care group [RR 1.45 (95% CI 0.79, 2.69)]. This body of evidence was downgraded for concerns regarding risk of bias and imprecision; a low quality GRADE rating was assigned.

B-2.2.3 LDCT Screening versus CXR Screening

A fourth RCT was a head-to-head trial comparing screening with LDCT with screening with CXR (NLST[10]). This study included 53,454 participants (26,722 LDCT; 26,732 CXR). The NLST study included mixed gender participants; about 60% of the sample was men. At enrollment, about 60% of participants were over age 60. Only current (48%) or former (quit in last 15 years) (52%) smokers were included. Participants in each group were offered three annual screens with their respectively assigned test. In terms of follow-up, all-cause mortality was assessed at a median of 6.5 years (maximum 7.4 years). This trial was initiated in 2002 in the US. The LDCT screening group showed a relative reduction of 6% in all-cause mortality as compared to the CXR screening group [RR 0.94 (95% CI 0.88, 1.00); absolute value per million 4,571 fewer, range from 180 fewer to 8,709 fewer]. The absolute risk reduction (ARR) is 0.46%. The number needed to screen (NNS) using LDCT to avert one death from any cause is 219 (95% CI 115, 5,556). This body of evidence was not downgraded on any domain; a high quality GRADE rating was assigned.

B-3.0 Lung Cancer Stage at Diagnosis

The point of cancer screening programs is to detect early stage disease in asymptomatic people that may be more responsive to treatment, in turn preventing progression to later stage disease and promoting longer survival. Therefore, if screening is working as intended, along with an expected increase in the number of diagnosed cases of early stage disease (with a potential risk for overdiagnosis) there should be a subsequent reduction in the number of people presenting with late stage disease in screened groups as compared to unscreened groups.

Eight studies were identified that provided data on lung cancer stage at diagnosis that could be pooled. Five RCTs used CXR as the primary screening test;[38, 39, 42, 43, 45] in four of these studies

CXR was combined with SC.[39, 42, 43, 45] Two RCTs used LDCT as the primary screening test.[21, 40] One study compared LDCT with CXR.[10] The results of these analyses are presented first for diagnoses of early stage disease and then for late stage disease. Findings from two additional studies[23, 41] could not be pooled and thus are reported narratively below. Our search and selection process did not locate any RCTs that investigated SC as the only screening test, that met our inclusion criteria and that reported on the outcome of stage at diagnosis.

Evidence Set 3 provides the forest plots (3.1 to 3.4) generated for the outcome of lung cancer stage at diagnosis. GRADE tables were not prepared for this outcome as it was rated as important (not critical) by the CTFPHC working group.

Early Stage I & II Non-Small Cell Lung Cancer (NSCLC)

B-3.1 Chest X-Ray (CXR)

Five RCTs that used CXR as the primary screening test reported results for stage at diagnosis, early stage I & II NSCLC;[38, 39, 42, 43, 45] these studies were separated into two groups for analysis.

B-3.1.1 CXR Screening versus Usual Care

A single RCT examined screening with CXR alone versus usual care (no formalized screening) (PLCO[38]). The PLCO trial enrolled 154,901 mixed gender (equally balanced) adults aged 55 to 74 years (about one-third in their 50s, half in their 60s). About half of the sample (52%) was identified as current or former smokers. Participants in the screening group were offered an annual CXR for 16 years as part of a comprehensive health check-up; the control group received usual care (screening was not offered or advised, but patients could be screened if they and/or their health care provider initiated testing). The PLCO began in 1993 and was conducted in the US. This study received a low risk of bias rating. A total of 2,821 cases of NSCLC were diagnosed; 37.33% of these diagnoses were for early stage (I & II) disease. The proportion of NSCLCs presenting as early stage (I & II) was significantly higher in the CXR screened group (39.7%) as compared to the usual care group (34.9%) [RR 1.14 (95% CI 1.03, 1.25)].

B-3.1.2 More Intensive CXR plus SC Screening versus Less Intensive Screening

The second group of studies included four RCTs that examined screening using CXR plus SC versus less intensive screening (CXR and SC but fewer screens and longer interval; CXR alone; advised to have annual CXR and SC) (Czech Study,[42] Johns Hopkins Study,[43] Mayo Lung Project,[39] Memorial Sloan-Kettering Study[45]). These four studies have a combined sample of 35,983 (17,983 more intensive screening; 18,000 less intensive screening). All four studies included only men. At enrollment, about one-quarter of the Mayo Lung Project sample was under age 50, another quarter was aged 50 to 54, two-fifths were aged 55 to 65, and the remaining 10% were aged 65 and older. In the Czech study about 38% of the men were enrolled in their 40s, about half (48%) were in their 50s, and 14% were in their early 60s. About one-third of the Memorial Sloan-Kettering sample was under age 50 when recruited, a bit more than half (54%) of the men were in their 50s, and the remaining quarter were mostly in their 60s, with a

small percentage (3%) in their 70s. In the Johns Hopkins study just under one-third (31%) of the men were in their late 40s, about a quarter (27%) were in their early 50s and another quarter (25%) were in their late 50s, and the remaining enrollees were mostly in their 60s. All four studies recruited only heavy smokers (current or recently quit). In the Mayo Lung Project 94% of the men smoked ≥1 pack/day, 97% had smoked for ≥20 years and 91% had ≥25 pack-years. Similarly, all men in the Memorial Sloan-Kettering study smoked ≥1 pack/day and 91% had ≥25 pack-years. Likewise all men in the Johns Hopkins study smoked ≥1 pack/day and 94% had ≥25 pack-years. Half of the men in the Czech study had a lifetime consumption of 150,000 to 250,000 cigarettes (approximately 20.55 to 34.25 pack-years) and the other half had smoked more. Two studies only enrolled participants who showed no signs of lung cancer on a preliminary screen. In the Mayo Lung Project, participants were offered CXR and SC every four months for a period of six years (up to 18 screens); the less intense screening arm had the eligibility CXR and SC test and was subsequently advised to get screened at least yearly, but these men were not offered systematic re-screening. The Czech study offered one group the dual screen every six months for three years (up to six screens) followed by three years of annual CXR testing; the less intense screening group received the dual test prior to randomization and again three years later, after which annual CXR was offered for three more years. The Memorial Sloan-Kettering and Johns Hopkins studies offered half of the men annual CXR and SC every four months for five years (initial recruits may have had an additional two or three years of screening), and the other half were offered annual screening with CXR alone for the same amount of time. All four studies were initiated in the 1970s. Three studies were conducted in the US and one in Czechoslovakia. Two of the studies were rated as having low risk of bias[43, 45] and the other two were rated as having unclear risk of bias.[39, 42] A combined total of 1,167 cases of NSCLC were diagnosed; 50.64% of these diagnoses were for early stage (I & II) disease. The proportion of NSCLCs presenting as early stage (I & II) was not significantly different in the group that received more intensive CXR plus SC screening (53.7%) as compared to the group that received less intensive screening (47.2%) [RR 1.15 (95% CI 0.98, 1.36) I^2=50%].

B-3.2 Low-Dose Computed Tomography (LDCT)

Three RCTs that used LDCT as the primary screening test reported results for stage at diagnosis, early stage I & II NSCLC;[10, 21, 40] these studies were separated into two groups for analysis.

B-3.2.1 Annual LDCT Screening versus Usual Care

The first group included two RCTs that examined annual screening with LDCT versus usual care (DANTE,[21] DLCST[40]). These studies have a combined sample of 6,576 (3,328 LDCT; 3,248 usual care). The DLCST trial included a mixed gender sample (55% men) while the DANTE study included only men. At enrollment, on average, the DLCST participants were in their late 50s (mean age 58 years) and the DANTE enrollees were slightly older (mean 64 years). Both studies included only current or former (quit in last 10 years) smokers. All participants in the DLCST study had ≥20 pack-years (mean 36 pack-years) and the mean number of cigarettes smoked per day was 19. In the DANTE study a little over half of the men were current smokers

and all the men also had ≥20 pack-years (mean 47 pack-years). Participants in the intervention arm of the DLCST trial were offered five annual LDCT screens; the control group received usual care. At baseline, men in the screening group of the DANTE trial had a CXR plus SC test in addition to their first of five annual LDCT scans; the usual care group also had the baseline CXR and SC test (no LDCT) but then only attended yearly for clinical reviews (no screening). Both studies were initiated in the last 10 to 15 years (since 2000). One was conducted in Italy and the other in Denmark. One study was rated as having unclear risk of bias[40] and the other was rated as having high risk of bias.[21] A combined total of 178 cases of NSCLC were diagnosed; 58.43% of these diagnoses were for early stage (I & II) disease. The proportion of NSCLCs presenting as early stage (I & II) was significantly higher in the annual LDCT screened group (65.9%) as compared to the usual care group (40.4%) [RR 1.59 (95% CI 1.11, 2.28) I^2=0%].

B-3.2.2 LDCT Screening versus CXR Screening

The second group included a single trial comparing screening with LDCT with screening with CXR (NLST[10]). The NLST study included 53,454 participants (26,722 LDCT; 26,732 CXR). The trial included mixed gender adults; about 60% of the sample was men. At enrollment, about 60% of participants were over age 60. Only current (48%) or former (quit in last 15 years) (52%) smokers were included. Participants in each group were offered three annual screens with their respectively assigned test. The NLST trial was initiated in 2002 in the US. This study received a low risk of bias rating. A total of 1,969 cases of NSCLC were diagnosed; 48.55% of these diagnoses were for early stage (I & II) disease. The proportion of NSCLCs presenting as early stage (I & II) was significantly higher in the annual LDCT screened group (57.0%) as compared to the CXR screened group (39.1%) [RR 1.46 (95% CI 1.33, 1.61)].

Late Stage III & IV Non-Small Cell Lung Cancer (NSCLC)

B-3.3 Chest X-Ray (CXR)

Five RCTs that used CXR as the primary screening test reported results for stage at diagnosis, late stage III & IV NSCLC;[38, 39, 42, 43, 45] these studies were separated into two groups for analysis.

B-3.3.1 CXR Screening versus Usual Care

A single RCT examined screening with CXR alone versus usual care (no formalized screening) (PLCO[38]). The PLCO trial enrolled 154,901 mixed gender (equally balanced) adults aged 55 to 74 years (about one-third in their 50s, half in their 60s). About half of the sample (52%) was identified as current or former smokers. Participants in the screening group were offered an annual CXR for 16 years as part of a comprehensive health check-up; the control group received usual care (screening was not offered or advised, but patients could be screened if they and/or their health care provider initiated testing). The PLCO began in 1993 and was conducted in the US. This study received a low risk of bias rating. A total of 2,821 cases of NSCLC were diagnosed; 62.67% of these diagnoses were for late stage (III & IV) disease. The proportion of NSCLCs presenting as late stage (III & IV) was significantly lower in the CXR screened group (60.3%) as compared to the usual care group (65.1%) [RR 0.93 (95% CI 0.87, 0.98)].

B-3.3.2 More Intensive CXR plus SC Screening versus Less Intensive Screening

The second group of studies included four RCTs that examined screening using CXR plus SC versus less intensive screening (CXR and SC but fewer screens and longer interval; CXR alone; advised to have annual CXR and SC) (Czech Study,[42] Johns Hopkins Study,[43] Mayo Lung Project,[39] Memorial Sloan-Kettering Study[45]). These four studies have a combined sample of 35,983 (17,983 more intensive screening; 18,000 less intensive screening). All four studies included only men. At enrollment, about one-quarter of the Mayo Lung Project sample was under age 50, another quarter was aged 50 to 54, two-fifths were aged 55 to 65, and the remaining 10% were aged 65 and older. In the Czech study about 38% of the men were enrolled in their 40s, about half (48%) were in their 50s, and 14% were in their early 60s. About one-third of the Memorial Sloan-Kettering sample was under age 50 when recruited, a bit more than half (54%) of the men were in their 50s, and the remaining quarter were mostly in their 60s, with a small percentage (3%) in their 70s. In the Johns Hopkins study just under one-third (31%) of the men were in their late 40s, about a quarter (27%) were in their early 50s and another quarter (25%) were in their late 50s, and the remaining enrollees were mostly in their 60s. All four studies recruited only heavy smokers (current or recently quit). In the Mayo Lung Project 94% of the men smoked ≥1 pack/day, 97% had smoked for ≥20 years and 91% had ≥25 pack-years. Similarly, all men in the Memorial Sloan-Kettering study smoked ≥1 pack/day and 91% had ≥25 pack-years. Likewise all men in the Johns Hopkins study smoked ≥1 pack/day and 94% had ≥25 pack-years. Half of the men in the Czech study had a lifetime consumption of 150,000 to 250,000 cigarettes (approximately 20.55 to 34.25 pack-years) and the other half had smoked more. Two studies only enrolled participants who showed no signs of lung cancer on a preliminary screen. In the Mayo Lung Project participants were offered CXR and SC every four months for a period of six years (up to 18 screens); the less intense screening arm had the eligibility CXR and SC test and was subsequently advised to get screened at least yearly, but these men were not offered systematic re-screening. The Czech study offered one group the dual screen every six months for three years (up to six screens) followed by three years of annual CXR testing; the less intense screening group received the dual test prior to randomization and again three years later, after which annual CXR was offered for three more years. The Memorial Sloan-Kettering and Johns Hopkins studies offered half of the men annual CXR and SC every four months for five years (initial recruits may have had an additional two or three years of screening), and the other half were offered annual screening with CXR alone for the same amount of time. All four studies were initiated in the 1970s. Three studies were conducted in the US and one in Czechoslovakia. Two of the studies were rated as having low risk of bias[43, 45] and the other two were rated as having unclear risk of bias.[39, 42] A combined total of 1,167 cases of NSCLC were diagnosed; 49.36% of these diagnoses were for late stage (III & IV) disease. The proportion of NSCLCs presenting as late stage (III & IV) was significantly lower in the group that received more intensive CXR plus SC screening (46.3%) as compared to the group that received less intensive screening (52.8%) [RR 0.85 (95% CI 0.75, 0.96) I^2=13%].

B-3.4 Low-Dose Computed Tomography (LDCT)

Three RCTs that used LDCT as the primary screening test reported results for stage at diagnosis, late stage III & IV NSCLC;[10, 21, 40] these studies were separated into two groups for analysis.

B-3.4.1 Annual LDCT Screening versus Usual Care

The first group included two RCTs that examined annual screening with LDCT versus usual care (DANTE,[21] DLCST[40]). These studies have a combined sample of 6,576 (3,328 LDCT; 3,248 usual care). The DLCST trial included a mixed gender sample (55% men) while the DANTE study included only men. At enrollment, on average, the DLCST participants were in their late 50s (mean age 58 years) and the DANTE enrollees were slightly older (mean 64 years). Both studies included only current or former (quit in last 10 years) smokers. All participants in the DLCST study had ≥20 pack-years (mean 36 pack-years) and the mean number of cigarettes smoked per day was 19. In the DANTE study a little over half of the men were current smokers and all the men also had ≥20 pack-years (mean 47 pack-years). Participants in the intervention arm of the DLCST trial were offered five annual LDCT screens; the control group received usual care. At baseline, men in the screening group of the DANTE trial had a CXR plus SC test in addition to their first of five annual LDCT scans; the usual care group also had the baseline CXR and SC test (no LDCT) but then only attended yearly for clinical reviews (no screening). Both studies were initiated in the last 10 to 15 years (since 2000). One was conducted in Italy and the other in Denmark. One study was rated as having unclear risk of bias[40] and the other was rated as having high risk of bias.[21] A combined total of 178 cases of NSCLC were diagnosed; 41.57% of these diagnoses were for late stage (III & IV) disease. The proportion of NSCLCs presenting as late stage (III & IV) was significantly lower in the annual LDCT screened group (34.1%) as compared to the usual care group (59.6%) [RR 0.59 (95% CI 0.43, 0.83) I^2=0%].

B-3.4.2 LDCT Screening versus CXR Screening

The second group included a single trial comparing screening with LDCT with screening with CXR (NLST[10]). The NLST study included 53,454 participants (26,722 LDCT; 26,732 CXR). The trial included mixed gender adults; about 60% of the sample was men. At enrollment, about 60% of participants were over age 60. Only current (48%) or former (quit in last 15 years) (52%) smokers were included. Participants in each group were offered three annual screens with their respectively assigned test. The NLST trial was initiated in 2002 in the US. This study received a low risk of bias rating. A total of 1,969 cases of NSCLC were diagnosed; 51.44% of these diagnoses were for late stage (III & IV) disease. The proportion of NSCLCs presenting as late stage (III & IV) was significantly lower in the annual LDCT screened group (43.0%) as compared to the CXR screened group (60.9%) [RR 0.71 (95% CI 0.65, 0.77)].

B-3.5 Narrative Results

The MILD study[23] included 4,099 participants (1,190 annual LDCT; 1,186 biennial LDCT; 1,723 usual care). The trial included mixed gender participants; about two-thirds of the sample was men. At enrollment, participants were mostly in their late 50s (median age 57 years). Only

current or former (quit in last 10 years) smokers were recruited; 90% of usual care and 69% of screening participants were current smokers, the overall mean pack-years was 38 and consumption for two-thirds of the sample was ≥20 cigarettes per day. In the intervention arms participants were offered LDCT screening either once a year (median of five screens) or every two years (median of three screens); the control group received usual care. Follow-up occurred at a median of 4.4 years (maximum six years). This trial, which was rated as having unclear risk of bias, was initiated in 2005 in Italy. Stage of diagnosis results could not be pooled with the other studies because no usual care group events were given. Study authors provided stage of diagnosis data only for the annual LDCT group and the biennial LDCT group. More than two-thirds of the cases across both groups (35/49) were detected at stage I or II (annual n=20, biennial n=15). There was no difference in the proportion of early or late stage lung cancers detected in annual versus biennial screening arms (P=0.53).

The Lung Screening Study[41] included 3,318 participants (1,658 CXR; 1,660 LDCT) who were recruited from PLCO screening sites but who were not taking part in the PLCO trial.[38] About 60% of the sample was men. The study targeted adults aged 55 to 74 years and at enrollment, about one-third of the participants were in their 50s, about one-third were in their early 60s, and the remaining third were aged 65 to 74 years. All participants were current (43%) or former (57%) smokers with ≥30 pack-years (half had <50 pack-years; half had ≥50 pack-years). This trial received an unclear risk of bias rating. Study authors reported stage of diagnosis results for each arm; however, to avoid biasing the results of a meta-analysis, the decision was made not to pool this data with the other trial (NLST[10]) comparing LDCT with CXR because of the less than adequate and incompatible follow-up period (≤12 months for the Lung Screening Study versus a median of 6.5 years for the NLST) and the small number of reported events. In the group that received LDCT scans, by the end of the study there were 19 diagnoses of early stage cancers, nine cases of late stage disease and two un-staged diagnoses (29/30 diagnoses were for NSCLC; one case of small cell carcinoma was detected but it is not clear at which stage it was diagnosed). Comparatively, in the CXR screening arm there were six diagnoses of early stage disease by the end of the study, no cases of late stage disease and one un-staged diagnosis (all seven cases were NSCLC).

B-4.0 Smoking Cessation Rate

A total of four studies reported data for smoking cessation rate: two studies used CXR combined with SC testing,[39, 42] and two studies used LDCT scans.[37, 40] Considering this self-reported outcome, all four studies received unclear risk of bias ratings. None of the data could be pooled; all results are reported narratively.

B-4.1 Chest X-Ray (CXR)

The Czech study,[42] which was initiated in 1976 in Czechoslovakia, included 6,364 male participants who showed no signs of lung cancer on a preliminary screening test (CXR plus SC). About 38% of the men were enrolled in their 40s, about half (48%) were in their 50s, and 14% were in their early 60s. Upon enrollment all participants were current and heavy smokers; half of the men had a lifetime consumption of 150,000 to 250,000 cigarettes (approximately 20.55 to

34.25 pack-years) and the other half had smoked more. Participants were randomized either to an intensive dual screening program (six CXR and SC tests at six month intervals) or to a less intense intervention involving annual CXR testing. After the initial three years, men in both study groups were offered annual CXR testing for an additional three years. There was no mention if, as part of the trial, participants were advised to quit smoking, if they received any smoking cessation counseling, and/or if they were given any printed materials about smoking cessation. At baseline there was no difference between study arms in terms of smoking history. In the intensive screening group the mean duration of smoking was 31.7 years and the mean lifetime consumption of cigarettes was 266,334; in the less intensive screening group the mean duration of smoking was and 31.5 years with an average lifetime consumption of 263,046 cigarettes. No specific data are reported; however study authors indicate that after six years of follow-up there was no difference between study groups in terms of the proportion who quit smoking or switched from cigarettes to pipes or cigars.[42]

The Mayo Lung Project,[39] which was initiated in the US in 1971, included 9,211 male participants who showed no signs of lung cancer on a preliminary screen. About one-quarter of the sample was under age 50, another quarter was aged 50 to 54, two-fifths were aged 55 to 65, and the remaining 10% were aged 65 and older. Only heavy smokers were recruited; 94% of the men smoked ≥1 pack/day, 97% had smoked for ≥20 years and 91% had ≥25 pack-years. In one arm participants were offered CXR and SC every four months for a period of six years (up to 18 screens); the less intense screening arm had the eligibility CXR and SC tests and was subsequently advised to get screened at least yearly, but these men were not offered systematic re-screening. No formal smoking cessation program was offered within the screening trial; however, during the initial interview with Lung Project staff (trained non-professionals supervised by a physician) and prior to group assignment, all participants were informed of the risks associated with smoking and the benefits of quitting, and all were advised to discontinue smoking. At the end of year one more than 90% of the men involved in the study either continued or resumed smoking (no information given whether this figure differed by intervention status). A 28-year follow-up survey asked participants (intensive screening n=537; less intensive screening n=549) or if deceased, their next-of-kin (intensive screening n=1,406; less intensive screening n=1,456) several questions about their smoking status. There were no statistically significant differences between study groups for any of these questions, regardless of whether answers were given by living participants or by their proxies. More than 80% of living men in both study arms reported they had not smoked in the previous 30 days and about 70% of these men indicated they had quit smoking more than 10 years earlier (before 1990). Roughly half of the living men who continued to smoke indicated that they were consuming less than one pack of cigarettes per day. Next-of-kin responses indicated slightly more than 60% of deceased men in both groups had not smoked in the year prior to their death; of those who had not quit, about one-third were smoking one to two packs per day and another third were smoking two or more packs per day.

B-4.2 Low-Dose Computed Tomography (LDCT)

The NELSON trial,[37] which was initiated in 2003 in the Netherlands and Belgium, included 15,822 mixed gender participants aged 55 to 75 years. Most enrollees were men (85.5%) and the average age was 58 years. Upon enrollment all participants were current (56%) or former (44%, quit <10 years prior) smokers. To be eligible participants had to have smoked ≥15 cigarettes per day for 25 years or ≥10 cigarettes per day for 30 years; the median pack-years was 38.0 (interquartile range 29.7 to 49.5). Participants were randomized either to four rounds of LDCT screening over 5.5 years or to a no screening control group. All current smokers, regardless of group status, were given either standard printed information about smoking cessation (e.g., barriers, advantages, relapse prevention, support options) or a tool (with questions about smoking history, attempts at abstinence, attitudes and self-efficacy regarding smoking cessation) to help tailor a request for information from the national tobacco control centre. After two years of participation, a subsample of men (641 LDCT; 643 control) was interviewed about their smoking status. There was no difference between study groups in terms of the number of attempts to stop smoking. Considering the data for responders only, a significantly higher point prevalence of smoking abstinence was observed in the control group compared to the screened group [OR 1.38 (95% CI 1.01, 1.90)], as well as lower prolonged and continued abstinence rates [OR 1.40 (95% CI 1.01, 1.92) and OR 1.42 (95% CI 1.03, 1.96) respectively]. However, when an ITT analysis was used (assuming non-responders as current smokers) there was no longer any significant difference between groups; the point prevalence of smoking abstinence in the screened group was 13.7% (n=88) compared with 15.5% (n=99) in the control group (P=0.38); continued smoking abstinence (<5 cigarettes since quit date) was 12.6% (n=81) in the LDCT group versus 14.6% (n=94) in the control group (P=0.30), and prolonged smoking abstinence (<5 cigarettes from two weeks after quit date) was 13.1% (n=84) versus 14.9% (n=96) in the control group (P=0.35).

The DLCST study,[40] which was initiated in Denmark in 2004, included 4,104 mixed gender participants (2,052 LDCT; 2,052 usual care); slightly more than half of the sample (55%) were men. At enrollment the mean age of participants was 58 years. The study recruited only current or former (quit after age 50 and in last 10 years) smokers with a history of ≥20 pack-years; enrollees had a mean of 36 pack-years and smoked a mean of 19 cigarettes per day. Participants in the intervention arm were offered five annual LDCT screens and the control group received usual care. Smokers in both groups received a brief session (<5 minutes) of smoking cessation counseling provided by trained nurses; former smokers were encouraged to continue abstinence. The study groups did not differ significantly in terms of annual smoking status. The number of LDCT group participants who self-identified as former smokers increased from 507 (25%) at baseline to 806 (43%) at the five year follow-up point. There was a comparable increase in the number of control group participants who quit smoking during the study; 473 (23%) self-identified as former smokers at baseline and 713 (43%) selected this category five years later.

B-5.0 Incidental Findings

A total of three studies reported data for incidental findings; one study used CXR combined with SC testing;[45] one used LDCT scans,[21] and the third compared LDCT with CXR.[10] No incidental findings were reported for SC testing alone. As mentioned in the USPSTF review,[26] there is little consistency in how incidental findings are reported which makes it difficult to meta-analyze.

B-5.1 Chest X-Ray (CXR)

The Memorial Sloan Kettering study,[45] which was initiated in the US in 1974, included 10,040 male participants (4,968 more intensive screening; 5,072 less intensive screening). At enrollment, about one-third of the sample was <50 years, a bit more than half (54%) of the men were in their 50s, and the remaining quarter were mostly in their 60s, with a small percentage (3%) in their 70s. Only current or former (time since quit ≤1 year) heavy smokers were included; all men smoked ≥1 pack/day and 91% had ≥25 pack-years. The study offered half of the men annual CXR and SC every four months for five years (initial recruits may have had an additional two or three years of screening), and the other half were offered annual screening with CXR alone for the same amount of time. Follow-up assessment was up to nine years in this low risk of bias study. In terms of incidental findings, a variety of non-neoplastic radiologic abnormalities were detected on initial CXRs including: emphysema, bullae, asbestosis, marked fibrosis, pleural effusion, active and old inflammatory disease, and abnormalities such as vascular, cardiac, mediastinal, skeletal, diaphragmatic and soft tissue. This study reported that these overall non-neoplastic radiologic abnormalities were detected in 50.9% of subjects receiving CXR alone and in 50.2% of those receiving CXR plus SC.

The NLST study,[10] which was initiated in the US in 2002, was a trial comparing screening with LDCT with screening with CXR. The study included 53,454 participants (26,722 LDCT; 26,732 CXR) and about 60% of this sample was men. At enrollment, approximately 60% of participants were over age 60 (targeted ages 55 to 74). Only current (48%) and former (52%, quit in last 15 years) smokers were included. Participants in each group were offered three annual screens with their respectively assigned test. Follow-up assessment was at a median of 6.5 years in this low risk of bias study. In terms of incidental findings, screening using the CXR test resulted in clinically significant findings other than lung cancer (no examples provided) in 2.1% of participants.[10]

B-5.1 Low-Dose Computed Tomography (LDCT)

The DANTE trial,[21] which was initiated in Italy in 2001, included 2,472 male participants (1,276 LDCT; 1,196 usual care). At enrollment, the mean age of the DANTE enrollees was 64 years (targeted 60 to 74 years). Only current (57%) and former (43%, quit in last 10 years) heavy smokers were included; all men had ≥20 pack-years (mean 47 pack-years). At baseline, men in the screening group of the DANTE trial had a CXR plus SC test in addition to their first of five annual LDCT scans; the usual care group also had the baseline CXR and SC test (no LDCT) but then only attended yearly for clinical reviews (no screening). This high risk of bias study had an average follow-up time of 2.8 years (range two to 79 months). The study authors reported that extra pulmonary abnormalities were detected in 37 of the 1,276 adults screened using LDCT:

effusions (n=4), pleural lesions (n=1), mediastinal masses (n=12), mediastinal lymph node enlargement (n=6), and other lesions such as hiatal hernias, aortic aneurysms, intra-thoracic goiters, renal masses, adrenal masses, and diaphragmatic paralysis (n=14).

In the LDCT arm of the NLST study,[10] screening using this test resulted in clinically significant findings other than lung cancer (no examples provided) in 7.9% of participants.

KQ1a. *What is the difference in screening effectiveness in populations and subgroups with varying risk for lung cancer (age, gender, smoking history)?*

Few included studies conducted or reported results for mortality sub-group analyses based on age, gender and/or smoking history.[38, 39, 42, 43, 46] Fewer still provided results of sub-group analyses that were conducted on the same data from the same observation point that we used for our overall analyses.[38, 39]

The PLCO trial,[38] which was initiated in the US in 1993, was the first major RCT of lung cancer screening to include women; the sample had about equal representation of men (49.5%) and women (50.5%). Adults aged 55 to 74 years were targeted for enrollment; about 33% of participants were in their 50s, 53% in their 60s, and 13% in their early 70s. The trial recruited smokers (52% current and former) as well as never smokers. Intervention group participants (n=77,445) were offered an annual CXR for 16 years as part of a comprehensive health check-up. The control group (n=77,456) received usual care (screening was not offered or advised, but patients could be screened if they and/or their health care provider initiated testing). The length of follow-up for lung cancer mortality was up to 13 years (median 11.9, mean 11.2, interquartile range 10 to 13). As observed in Evidence Set 1, Forest Plot 1.1, there was no difference in lung cancer mortality between the CXR and usual care groups. The study authors reported similar non-significant results for this outcome when screened participants were compared to unscreened participants using sub-groups based on smoking status [never smokers RR 0.94 (95% CI 0.69, 1.29), former smokers RR 1.02 (95% CI 0.91, 1.15), current smokers RR 0.99 (95% CI 0.88, 1.12)] and gender [men RR 1.02 (95% CI 0.92, 1.13), women RR 0.92 (95% CI 0.81, 1.06)].

The Mayo Lung Project,[39] which was initiated in the US in 1971, included 9,211 male participants who showed no signs of lung cancer on a preliminary screen. About one-quarter of the sample was under age 50, another quarter was aged 50 to 54, two-fifths were aged 55 to 65, and the remaining 10% were aged 65 and older. Only heavy smokers were recruited; 94% of the men smoked ≥1 pack/day, 97% had smoked for ≥20 years and 91% had ≥25 pack-years. In one arm participants were offered CXR and SC every four months for a period of six years (up to 18 screens); the less intense screening arm had the eligibility CXR and SC tests and was subsequently advised to get screened at least yearly, but these men were not offered systematic re-screening. In terms of follow-up, lung cancer mortality was assessed at a median of 20.5 years. Adjusting for key risk factors for lung cancer (including age and pack-years) did not change the non-significant difference in disease specific mortality between the intensive and less intensive screening groups [unadjusted hazard ratio (HR) 1.1 (95% CI 1.0, 1.3); adjusted HR 1.1

(95% CI 1.0, 1.3)]. Sub-group analyses for age and pack-years also showed non-significant differences between study arms [aged <55 HR 1.0 (no CIs reported), aged 55 to 64 HR 1.1, aged ≥65 HR 1.6; <50 pack-years HR 1.1, 50 to 99 pack-years HR 1.1, ≥100 pack-years HR 1.1].

KQ2. *What are the harms (H) of screening for lung cancer in adults not suspected of having lung cancer?*

The harms of interest for this review include: overdiagnosis, death from invasive follow-up testing, major complications or morbidity from invasive follow-up testing, false positives, consequences of false positives (i.e., people with benign conditions undergoing invasive follow-up procedures), negative consequences of incidental findings, anxiety, quality of life, infection from invasive follow-up testing, and bleeding from invasive follow-up testing. Our search located 30 studies that provided evidence of one or more of these harms from lung cancer screening using LDCT, CXR and/or SC tests or from invasive follow-up procedures. Where possible, data are presented for the harms by number of tests (e.g., invasive procedures) and/or number of patients. Most of the harms data was obtained from observational studies; using the GRADE approach the overall quality of the evidence was rated as LOW.

Key findings across critical outcomes (i.e., overdiagnosis, death and major complications from invasive follow-up testing), with pooled estimates of effect, are provided in Table 6. Detailed results for critical and important (i.e., false positives, consequences of false positives, negative consequences of incidental findings, anxiety, quality of life, infection or bleeding from invasive follow-up testing) outcomes are presented below.

H-1.0 Overdiagnosis

In this review, overdiagnosis refers to the detection of a lung cancer that will not otherwise cause symptoms throughout the person's lifetime or result in death. Five papers about six trials were identified that provided data on overdiagnosis.[52, 65-68] Evidence Set 4 provides the Findings Summary and GRADE Rating Tables (ES Tables 4.1 and 4.2) generated for the outcome of overdiagnosis.

H-1.1 Chest X-Ray (CXR) plus Sputum Cytology (SC)

One paper[66] provided data for two clinical trials which used CXR plus SC as the screening intervention. The Mayo Lung Project offered participants a dual screening test using CXR and SC every four months. The Memorial Sloan-Kettering study offered participants annual CXR screening with the addition of SC tests repeated every four months. The percentage of overdiagnosis was reported for the cut-points of tumor volume doubling time (TVDT) >400 and >300 days. For the cut-point of TVDT >400 days, of all cases of lung cancer diagnosed in the screened population an estimated 2.27% to 6.98% were overdiagnosed. For the cut-point of TVDT >300 days, overdiagnosis estimates ranged from 4.55% to 16.28%. The overall quality of this body of evidence was rated as LOW.

H-1.2 Low-Dose Computed Tomography (LDCT)

Four studies provided overdiagnosis data for LDCT as the screening test. In the United Kingdom Lung Screening (UKLS) study participants in the intervention arms received either annual or biennial LDCT scans. Annual scans were offered to participants in one arm of the NLST trial, to participants in an observational study conducted in rural Japan, and to participants in the COSMOS trial. The four papers providing overdiagnosis data for these trials[52, 65, 67, 68] used varied and unclear cut-points (lead time 5.5 years with mean sojourn time two years, lead time ≥5 years, tumour size 30 mm, TVDT ≥400 days), and for the NLST study separate results were reported for screen detected and diagnosed cancers. Across these conditions, of all cases of lung cancer diagnosed in the screened population an estimated 10.99% to 25.83% were overdiagnosed. The overall quality of this body of evidence was rated as LOW.

H-2.0 Death from Invasive Follow-up Testing

Death from follow-up testing refers to mortality that is the direct consequence of an invasive follow-up procedure (e.g., video-assisted thoracoscopic surgery, fine-needle aspiration biopsy or fine-needle aspiration cytology, thoracotomy, bronchoscopy, mediastinoscopy, surgical resection) initiated as a result of screening. Twelve papers were identified that provided data on death from invasive follow-up testing across a variety of screening trials.[10, 21, 50, 53, 54, 57, 63, 64, 69-72] Evidence Set 5 provides the GRADE Rating Table (ES Table 5.1) and the forest plot (5.1) generated for this outcome. All bodies of evidence received low GRADE ratings due to the observational nature of the included studies, the variation observed across types of procedures and length of follow-up, and in some papers the reporting of data only for patients with lung cancer.

H-2.1 Chest X-Ray (CXR)

Deaths resulting from invasive procedures that followed CXR screening tests were reported for five screening arms across four studies.[10, 50, 64, 69] Out of 778 patients who underwent invasive follow-up procedures, 23 died; resulting in an absolute number of 28.60 deaths (95% CI 16.02, 41.17) per 1,000 patients undergoing invasive follow-up testing.

H-2.2 Chest X-Ray (CXR) plus Sputum Cytology (SC)

Deaths resulting from invasive procedures that followed dual CXR plus SC screening were reported for four screening arms across three studies.[69-71] Out of 333 patients who underwent invasive follow-up procedures, 21 died; resulting in an absolute number of 47.67 deaths (95% CI 23.86, 71.49) per 1,000 patients undergoing invasive follow-up testing.

H-2.3 Low-Dose Computed Tomography (LDCT)

Deaths resulting from invasive procedures that followed LDCT scans were reported for screening arms across seven studies.[10, 21, 53, 54, 57, 63, 72] Out of 1,502 patients who underwent invasive follow-up procedures, 20 died; resulting in an absolute number of 11.18 deaths (95% CI 5.07, 17.28) per 1,000 patients undergoing invasive follow-up testing.

H-3.0 Major Complications or Morbidity from Invasive Follow-up Testing

This section considers major complications or morbidity requiring hospitalization or medical intervention (e.g., hemothorax and pneumothorax requiring tube placement, lung collapse, severe pain, cardiac arrhythmias and thromboembolic complications[73, 74]) that are the direct result of an invasive procedure (e.g., video-assisted thoracoscopic surgery, fine-needle aspiration biopsy or fine-needle aspiration cytology, thoracotomy, bronchoscopy, mediastinoscopy, surgical resection) initiated as a result of screening. Four papers were identified that provided data on major complications or morbidities from invasive follow-up testing across several screening trials.[10, 54, 57, 72] Evidence Set 6 provides the GRADE Rating Table (ES Table 6.1) and the forest plot (6.1) generated for this outcome. Both bodies of evidence received low GRADE ratings due to the observational nature of the included studies, the variation observed across types of procedures and length of follow-up, and in some papers the reporting of data only for patients with lung cancer.

H-3.1 Chest X-Ray (CXR)

Major complications or morbidity resulting from invasive procedures that followed CXR screening tests were reported for one study.[10] Out of 379 patients who underwent invasive follow-up procedures, 24 had major complications or morbidity; resulting in an absolute number of 63.32 major complications (95% CI 42.92, 92.49) per 1,000 patients undergoing invasive follow-up procedures.

H-3.2 Low-Dose Computed Tomography (LDCT)

Major complications or morbidity resulting from invasive procedures that followed LDCT scans were reported for screening arms across four studies.[10, 54, 57, 72] Out of 1,336 patients who underwent invasive follow-up procedures, 92 had major complications or morbidity; resulting in an absolute number of 43.29 major complications (95% CI 32.00, 54.58) per 1,000 patients undergoing invasive follow-up procedures.

H-4.0 False Positives

In this review, a false positive refers to a screening test result that indicates the presence of lung cancer, when in fact no lung malignancy exists. Nine papers were identified that provided false positive data across a variety of RCTs and observational screening studies.[10, 19, 24, 49-51, 56, 75, 76] Evidence Set 7 provides the Findings Summary Table (ES Table 7.1) generated for this outcome.

H-4.1 Chest X-Ray (CXR)

False positive data were extracted for three studies that offered participants multiple rounds of CXR screening.[10, 50, 51] The cut-point for a positive test varied across studies; two studies used a mass >3 cm or a nodule <3 cm and the third study used a suspicious nodule. Out of the 33,199 subjects who underwent screening in these trials, 2,098 received at least one false positive result [median 6.50% (range 3.40% to 13.67%)]; resulting in a median absolute number of 65.0 false positives (range 34.0 to 136.7) per 1,000 screened.

H-4.2 Low-Dose Computed Tomography (LDCT)

False positive data were extracted for the baseline tests conducted in two studies using LDCT.[19, 49] The cut-point for a positive test varied; one study used a nodule >5 mm and the other study used a nodule >10 mm. Out of the 4,081 subjects who underwent the baseline LDCT screening in these trials, 680 received at least one false positive result [median 16.71% (range 7.90% to 25.53%)]; resulting in a median absolute number of 167.1 false positives (range 79.0 to 255.3) per 1,000 screened.

False positive data were also extracted for seven studies that offered participants multiple rounds of LDCT screening.[10, 19, 24, 51, 56, 75, 76] The cut-point for a positive test varied across studies, ranging from >3 mm nodules to >8 mm nodules, and two studies set different cut-points for solid and non-solid nodules. Out of the 42,774 subjects who underwent LDCT screening in these trials, 8,290 received at least one false positive result [median 23.30% (range 0.64% to 69.0%)]; resulting in a median absolute number of 233.0 false positives (range 6.4 to 690.0) per 1,000 screened.

H-5.0 Consequences of False Positives

This section considers the consequences of false positives, specifically patients with benign conditions undergoing minor (e.g., fine-needle aspiration biopsy or fine-needle aspiration cytology, thoracic or lymph node biopsy, bronchoscopy) or major (e.g., video-assisted thoracoscopic surgery, thoracotomy, surgical resection) invasive procedures initiated as a result of false positive screening tests. Nine papers were identified that provided data on patients undergoing minor invasive procedures as a consequence of false positives[41, 48, 50, 51, 58, 61, 62, 75, 77] and 19 papers were found that included data for patients undergoing major invasive procedures as a consequence of false positives.[19, 21, 23, 24, 41, 48, 50, 51, 53, 54, 57-61, 63, 75, 77, 78] Evidence Set 8 provides the Findings Summary Table (ES Table 8.1) and forest plots (8.1 and 8.2) generated for this outcome.

H-5.1 Minor Invasive Procedures - Chest X-Ray (CXR)

Minor invasive procedures that followed false positive CXR screening results were reported for four observational studies.[41, 50, 51, 77] Out of 81,819 people who underwent screening, 192 individuals with benign conditions were subjected to minor invasive procedures as part of diagnostic follow-up; resulting in an absolute number of 2.30 subjects with benign conditions undergoing minor invasive procedures (95% CI 1.49, 3.11) per 1,000 screened.

H-5.2 Minor Invasive Procedures - Low-Dose Computed Tomography (LDCT)

Minor invasive procedures that followed false positive LDCT screening results were reported for seven observational studies.[41, 48, 51, 58, 61, 62, 75] Out of 15,101 people who underwent screening, 81 individuals with benign conditions were subjected to minor invasive procedures as part of diagnostic follow-up; resulting in an absolute number of 7.16 subjects with benign conditions undergoing minor invasive procedures (95% CI 3.27, 11.05) per 1,000 screened.

H-5.3 Major Invasive Procedures - Chest X-Ray (CXR)

Major invasive procedures that followed false positive CXR screening results were reported for four observational studies.[41, 50, 51, 77] Out of 81,819 individuals who underwent screening, 139 patients with benign conditions were subjected to major invasive procedures as part of diagnostic follow-up; resulting in an absolute number of 2.73 subjects with benign conditions undergoing major invasive procedures (95% CI 0.96, 4.51) per 1,000 screened.

H-5.4 Major Invasive Procedures - Low-Dose Computed Tomography (LDCT)

Major invasive procedures that followed false positive LDCT screening results were reported for 17 observational studies.[19, 21, 23, 24, 41, 48, 51, 53, 54, 57-61, 63, 75, 78] Out of 41,411 individuals who underwent screening, 228 patients with benign conditions were subjected to major invasive procedures as part of diagnostic follow-up; resulting in an absolute number of 4.98 subjects with benign conditions undergoing major invasive procedures (95% CI 3.68, 6.29) per 1,000 screened.

H-6.0 Negative Consequences of Incidental Findings

As noted above in section B-5.0 of the results for key question one, a total of three studies reported data for incidental findings; one study used CXR combined with SC testing;[45] one used LDCT scans,[21] and the third compared LDCT with CXR.[10] We found no evidence on the consequences of incidental findings; this is not surprising given that these data are likely reported in treatment papers for those specific conditions, all of which were outside the scope of this systematic review.

H-7.0 Anxiety

Two studies that used LDCT scans as the test reported data for anxiety assessed in lung cancer screening participants.[55, 79] The data could not be pooled; results are presented narratively below.

H-7.1 Low-Dose Computed Tomography (LDCT)

Before and after screening using LDCT scans, 600 participants in the NELSON trial completed the short form of the State-Trait Anxiety Inventory. The overall score on this six item test can range from 20 to 80; higher scores indicate more anxiety. The baseline mean score was 33.2 (SD 8.6) and the post intervention score was 33.00 (SD 9.2); both scores fall within the low anxiety range (20 to 39). The computed mean change score (MCS) showed no significant difference in the summary score (i.e., participants' level of overall anxiety) over time [-0.20 (95% CI -0.65, 0.25)].

Before and after receiving lung cancer screening using LDCT scans, 393 participants in the PLuSS trial completed the short form of the State-Trait Anxiety Inventory. The overall score on this six item test can range from 20 to 80; higher scores indicate more anxiety. Study authors reported results for state (at the moment) anxiety separate from trait (general) anxiety. For state anxiety the baseline mean score was 35.16 (SD 12.36) and the post intervention mean score was 36.69 (SD 12.99); both scores fall within the low anxiety range (20 to 39). The computed MCS showed a significant increase in the summary score (i.e., participants' level of state anxiety) over time [1.53 (95% CI 0.74, 2.33)]. For trait anxiety the baseline mean score was 36.70 (SD 11.36) and the post intervention mean score was 36.92 (SD 11.60); both scores fall within the low

anxiety range (20 to 39). The computed MCS showed no difference in the summary score (i.e., participants' level of trait anxiety) over time [0.22 (95% CI -0.50, 0.94)].

H-8.0 Quality of Life

Data for health related quality of life were reported for three trials.[47, 79, 80] Due to variations in measures, data from two studies were plotted but not pooled (Evidence Set 9, forest plot 9.1). Only a within group analysis was possible for the third study; results are presented narratively below.

H-8.1 Chest X-Ray (CXR)

One trial observed no difference between screened (n=711) and control group (n=713) participants in terms of change in overall health related quality of life measured at baseline (pre) and post screening using the EuroQOL five dimensions questionnaire (EQ-5D) [mean difference (MD) 0.036 (95% CI 0.030, 0.043)].[47] Analyses of data provided by a sub-study of the PLCO trial[80] that administered the 12-item Short-Form questionnaire (SF-12) to 149 participants in the screening arm and 179 people in the usual care group found no difference between groups in terms of change from baseline to post-screening on the physical health composite score [MD -0.28 (95% CI -1.48, 0.93)], but found a significant difference between groups in terms of change on the mental health composite score, in favour of those who did not undergo screening [MD 1.19 (95% CI 0.20, 2.17)].

H-8.2 Low-Dose Computed Tomography (LDCT)

Before and after receiving lung cancer screening using LDCT scans, 600 participants in the NELSON trial completed the EQ-5D and the SF-12 to assess three dimensions of health related quality of life: overall, physical and mental.[79] The EQ-5D uses a visual analogue (thermometer-style) scale to indicate general health status assigning 100 as the best score possible. The SF-12 also uses a scoring range from zero to 100, with better health indicated by higher scores. For overall quality of life the baseline (pre-screening) mean score was 79.3 (SD 13.7) and the post-screening mean score was 78.4 (SD 13.7); both scores fall within the upper range. The computed MCS showed a significant decrease in the summary score (i.e., participants' overall health related quality of life) over time [-0.90 (95% CI -1.59, -0.21)]. For the physical dimension of quality of life, the baseline mean score was 49.50 (SD 8.7) and the post intervention mean score was 50.0 (SD 8.2); both scores fall within the normal range. The computed MCS showed a significant improvement in the summary score (i.e., participants' physical quality of life) over time [0.50 (95% CI 0.07, 0.93)]. Finally, for the mental dimension of quality of life, the baseline mean score was 51.9 (SD 10.3) and the post intervention mean score was 51.6 (SD 11.1); again, both scores fall within the normal range. The computed MCS showed no difference in the summary score (i.e., participants' mental quality of life) over time [-0.30 (95% CI -8.84, 0.24)].

H-9.0 Infection from Invasive Follow-up Testing

In this review, infection from follow-up testing refers to infection that is the direct consequence of an invasive procedure (e.g., video-assisted thoracoscopic surgery, fine-needle aspiration

biopsy or fine-needle aspiration cytology, thoracotomy, bronchoscopy, mediastinoscopy, surgical resection) initiated as a result of lung cancer screening. Two papers were identified that provided data on infection from invasive follow-up testing.[41, 81] Evidence Set 10 provides the Findings Summary Table (ES Table 10.1) and the forest plot (10.1) generated for this outcome.

H-10.1 Chest X-Ray (CXR)

Infections resulting from procedures that followed CXR screening were reported in two studies.[41, 81] Out of 621 patients who underwent invasive follow-up procedures, 35 developed an infection; resulting in an absolute number of 52.75 cases of infection (95% CI 35.10, 70.40) per 1,000 patients undergoing invasive follow-up testing.

H-10.2 Low-Dose Computed Tomography (LDCT)

Infections resulting from procedures that followed LDCT screening were reported in one study.[41] Out of 53 patients who underwent invasive follow-up procedures, two developed an infection; resulting in an absolute number of 37.74 cases of infection (95% CI 10.41, 127.54) per 1,000 patients undergoing invasive follow-up procedures.

H-10.0 Bleeding from Invasive Follow-up Testing

No studies were found that reported cases of bleeding that were the direct consequence of invasive procedures (e.g., video-assisted thoracoscopic surgery, fine-needle aspiration biopsy or fine-needle aspiration cytology, thoracotomy, bronchoscopy, mediastinoscopy, surgical resection) initiated as a result of lung cancer screening. Therefore no eligible evidence was available to address this outcome.

KQ2a. *What is the difference in harms in populations and subgroups with varying risk for lung cancer (age, gender, smoking history)?*

None of the included harms studies conducted or reported results for overdiagnosis, mortality or major complications sub-group analyses based on age, gender and/or smoking history. Consequently, there is no evidence on which to base a response to this question.

Results for Contextual Questions

CQ1. *What is the evidence that test characteristics for effective lung cancer screening tests (sensitivity and specificity, false positives and false negatives, negative and positive predictive values, and test positivity rate) differ by subgroups with varying risk for lung cancer?*

For this question we included one systematic review[26] and data from one RCT.[82] The USPSTF review[26] reported data from two RCTs and five cohort studies. The USPSTF reported that LDCT showed high sensitivity with a range of 80% to 100% in incidence and prevalence screens with most studies showing sensitivity greater than 90%. Analysis from the NLST, which reported on two rounds of screening, showed sensitivity at T1 (first screen) of 94.4% (95% CI 90.8, 97.6)

and 93.0% (95% CI 89.7, 96.3) at T2 (second screen). Specificity of LDCT reported by the USPSTF was a range of 28% to 100%. The specificity reported in the NLST trial was 72.6% (95% CI 72.0, 73.1) at T1 and 83.9% (95% CI 83.4, 84.3) at T2. The method of determining sensitivity and specificity varied across studies as there is no defined "gold or reference" standard for LDCT. The USPSTF reported a calculated positive predictive value (PPV) for an abnormal LDCT scan predicting lung cancer that ranged from 2.2% to 42%. A PPV of 2.4% (95% CI 2.1, 2.8) at T1 and 5.2% (95% CI 4.6, 5.9) at T2 was reported by the NLST. Aberle et al.[82] also provided a negative predictive value (NPV) of the NLST trial of 99.9% (95% CI 99.9, 100.0) at time T1 and T2.

CQ2. *What is the difference in test performance with changes and improvements in low-dose computed tomography technology or varying protocols used by radiologists?*

None of the included studies for test properties directly compared various LDCT technologies or varying protocols used by radiologists and their effect on test performance. To answer this question we looked at the studies included for the contextual question above for information on: use of single or multi detectors, slice width (mm), computer assisted reading/diagnosis (CAR/D), number of readers, and nodule size as a cut-point or threshold for a positive test. Four papers provided relevant data.[82-85] Three study protocols reported multi or variable detectors;[83-85] one study did not report on detectors.[82] One study used CAR/D by two radiologists[83] while the other studies did not include CAR/D with two reviewers;[84, 85] the range for sensitivity was 88.9% (95% CI 79.7, 98.1) to 100.0% (CI not reported) and the range for specificity was 92.6% (95% CI 92.0, 93.2) to 96.9% (CI not reported). One study[82] that used only one reader and did not use CAR/D reported sensitivity of 94.4% and specificity of 72.6%. The diagnostic test properties were highest overall in the study[83] that used a multi slice detector along with CAR/D plus two independent radiologist readers showing sensitivity of 94.6% (95% CI 86.5, 98.6), specificity of 98.3% (95% CI 98.0, 98.6), PPV of 35.7% (95% CI 29.3, 42.7), and NPV of 99.9% (95% CI 99.9, 100.0).

CQ3. *What are participants' values and preferences on screening for lung cancer?*

Seven papers (five uncontrolled observational studies and two qualitative studies) were found that addressed the question of participants' values and preferences regarding lung cancer screening.[86-92] This section includes studies that considered participants' willingness to be screened, motivators and barriers to attend screening, and participants' experiences with screening.

Willingness or Intent to be Screened and/or Treated

Four uncontrolled observational studies examined high risk individuals' willingness or intent to be screened for lung cancer.[86, 87, 91, 92]

To fill a gap in the literature about chronic obstructive pulmonary disease (COPD) patients' attitudes toward lung cancer screening and subsequent treatment of screen-detected disease, a 2012 study compared data from a mixed gender sample of 142 Irish adults with COPD and a

high risk smoking history to previously published data from an American general population-based sample of 1,076 current and former smokers.[86] Results showed that 97.9% of the Irish COPD sample was willing to undergo LDCT screening compared to 78.6% of the US group. A higher proportion of the COPD adults believed they were at risk for lung cancer (63.6% versus 15.7%) and that early detection could improve survival (90% versus 51.2%). The study also found that, compared to the US sample, the Irish COPD adults were more willing to consider paying $200/€200 for a screening test (68.6% versus 36.2%) and undergo treatment if it was recommended (95.7% versus 56.2%).

A 2012 US study examined racial and ethnic differences in willingness to be screened for lung cancer in participants with a ≥10 pack-year cigarette smoking history and found intentions were similar across the groups (76% non-minority, 90% Black, 77% Hispanic; P=0.19).[87] Factors that influenced participants' intentions to be screened (including cost, fatalistic beliefs, fear and anxiety) are discussed below.

In a study on attitudes and beliefs toward lung cancer screening of current, former, and never smoker US veterans, 92.8% of those surveyed, regardless of smoking status, indicated they would undergo a CT scan for lung cancer screening.[91] The study also reported that 92.4% would undergo surgery if recommended.

In the survey responses of 90 high risk, current and former smokers in Australia, a high level of willingness to participate in screening was reported (86/90).[92] The four participants unwilling to undergo lung cancer screening were current smokers.

Reasons to Participate in Screening

Six studies were found that reported on factors that motivate individuals to undergo lung cancer screening.[86, 88-92]

Smoking History

Two studies (one uncontrolled observational and one qualitative) found that smoking history or smoking status was a reason to participate in lung cancer screening.[89, 90]

In an uncontrolled observational study of a subsample of adults who participated in the NELSON trial, the reasons to participate in lung cancer screening among high risk individuals were examined. Results indicated that 79.2% of the 889 individuals who had participated in screening reported "I smoked a lot" as one of the reasons for getting tested, with 25.2% listing this as the decisive reason to participate.[89] Similarly, a qualitative study found most participants in a sample of 35 current and former smokers with a history of ≥30 pack-years recruited from four American College of Radiology Imaging Network sites, cited "heavy" smoking history as a reason to undergo screening.[90]

Early Detection

Another key reason to seek lung cancer screening was the belief that early detection of disease could lead to improved outcomes in terms of treatment and survival.[86, 88, 89, 91, 92]

In a sub-study of the NELSON trial, 79% of 889 participants who underwent screening reported early detection of lung cancer as one of the reasons for participating with 34.2% citing this as the decisive reason they underwent screening.[89] Likewise, a qualitative study involving interviews with 60 current or former smokers with a ≥20 pack-year smoking history and/or ≥ 20 years of smoking in the UK-based Lung-SEARCH trial reported that early detection was the "main hoped-for personal benefit" as a result of undergoing screening.[88]

Three additional studies reported that some participants believed early detection could improve chances of survival: 50% of participants in a survey of 209 current, former and never smoker veterans held this belief;[91] 72.2% of 90 high risk current and former smokers believed chances of survival would be "somewhat higher" or "very much higher" if cancer was detected early;[92] and 90% of 142 Irish COPD smokers also believed that early detection could improve chances of survival.[86] The authors of these three studies do not state that this belief motivated participants to undergo screening. However, each of these studies found a high willingness to participate in screening among their populations (92.8%, 95.5%, and 97.9%, respectively).

Familiarity with Lung Cancer

Two studies reported a familiarity with lung cancer (knowing someone who had lung cancer or family history) as a motivator for screening.[88, 90]

In a sub-study of the NELSON trial, 19% of participants stated family history was one of their motivating factors and 10.6% stated a history of lung cancer in acquaintances as a motivating factor to participate in lung cancer screening.[89]

In the Lung-SEARCH study, family history of lung cancer was also a motivating factor among some participants providing annual sputum samples, as well as those participants receiving annual bronchoscopy and CT scanning.[88] Conversely, the group of individuals who declined to participate in either type of screening viewed a lack of family history as a type of protection "against the effects of their continued smoking."

Risk Perception

Four studies linked perceived risk of lung cancer with the decision to participate in lung cancer screening.[86, 88, 89, 91]

Sixty interviews were conducted with current or former smokers with a ≥20 pack-year smoking history and/or ≥20 years of smoking in who either agreed to participate in the UK-based Lung-SEARCH trial (n=36) or who declined to take part in the screening (n=24). Three-quarters (28/36) of those who took part in the trial (sputum or CT scan plus bronchoscopy) perceived themselves to be at risk for lung cancer which informed their decision to participate.[88] In the group of 24 individuals who did not participate in screening there was underestimation and denial of risk. Similarly, in a sub-study of the NELSON trial, significantly more, 14.4%, of the participants who underwent screening (n=889) reported a high or very high belief that they would develop lung cancer as compared to only 6.5% of those who did not participate in the trial (n=79) (P=0.049).[89]

In a 2012 study comparing a mixed gender sample of 142 Irish adults with COPD and a high risk smoking history to an American general population-based sample of 1,076 current and former smokers, 83.1% vs. 68.1% believed risk of disease was an important factor in making a decision to be screened.[86] In another study on attitudes and beliefs of current, former, and never smoker US veterans towards lung cancer screening, 84.7% of all participants also believed that risk of disease was an important factor in their decision to be screened [OR 2.2 (95% CI 0.8, 6.0) for smokers vs. never smokers].[91]

One additional study found high risk perceptions among a sample of 35 current and former smokers with a history of ≥30 pack-years, but the authors reported that these perceptions did not appear to motivate individuals to undergo lung cancer screening.[90]

Barriers to Participating in Screening

Five studies reported on barriers to uptake of lung cancer screening.[87-89, 91, 92] Contextual question six below provides more details regarding barriers to screening in particular sub-populations.

Perceptions of Health Care Workers and Hospitals

Mistrust in health care workers and/or prior negative experiences with hospitals were stated as barriers to lung cancer screening in two studies.[87, 88]

Qualitative interviews were conducted with 60 mixed gender current or former smokers with a ≥20 pack-year smoking history and/or ≥20 years of smoking in the UK-based Lung-SEARCH trial. Negative experiences with hospitals and doctors were cited by some as a reason for not getting screened.[88]

A US-based uncontrolled observational study of 108 mixed gender and ethnically diverse participants (37% Black; 32% Hispanic and 32% non-minority) found that reluctance to get screened for lung cancer among those surveyed was significantly associated with a "mistrust in health care workers" (P=0.01).[87] The study reported that 3% of non-minorities, 5% of Blacks and 15% of Hispanics agreed with the statement "my ethnic group cannot trust doctors and health care workers."

Travel and Convenience of Screening

Two studies listed "screening convenience" as an important factor in decision making about whether to undergo lung cancer screening.[88, 91]

In a US study on attitudes and beliefs of current, former and never smoker US veterans towards lung cancer screening, 62.2% of all participants believed screening convenience was an important factor in making the decision about whether to be screened.[91]

Of the 24 interviewees who chose not to participate in annual sputum or bronchoscopy plus CT scan in the UK-based Lung-SEARCH trial, half cited travel as the most significant factor influencing their decision not to participate in the screening study.[88]

Other Barriers

In two studies that examined individuals' reasons for declining participation in a screening trial, the following were found to be deterrents or barriers to screening: negative perception of bronchoscopy (18/24),[88] and participation in a trial perceived as requiring too much effort (45.2% of 79 non-participants).[89] In a US-based uncontrolled observational study that surveyed 108 mixed gender and multi-ethnic participants (37% Black; 32% Hispanic and 32% non-minority), holding fatalistic beliefs about lung cancer (beliefs that reasons for lung cancer are not important and should just be accepted, or that it is better not to find out) was associated with a reluctance to get screened (P=0.01).[87] Finally, although an accurate risk perception may be a reason for some to consider lung cancer screening, an inaccurate risk perception may act as a barrier to screening. An uncontrolled observational study of high risk, current and former smokers in Australia reported that "under-recognition of risk" could be a barrier to lung cancer screening.[92]

Experience based Attitudes and Perceptions of Lung Cancer Screening Tests

Three studies reported on screening-experienced individuals' attitudes and perceptions of lung cancer screening.[88-90]

A qualitative study found most participants in a sample of 35 current and former smokers with a history of 30 pack-years did not describe lung cancer screening (low-dose computed tomography or chest x-ray) experiences as "stressful;" instead participants reported feeling the procedure was "just routine" and that after screening they felt "fine."[90]

Another qualitative study of 60 current or former smokers with a ≥20 pack-year smoking history and/or ≥20 years of smoking in the UK-based Lung-SEARCH trial found that screening methods, including annual sputum sample, bronchoscopy and CT scans were considered broadly "acceptable" by those who participated; 59/60 had positive views of CT scans and most participants were not concerned about providing annual sputum samples.[88]

In an uncontrolled observational study of a subsample of adults who participated in the NELSON trial, those who participated in CT screening had a more positive attitude toward lung cancer screening (98.7%) than those who declined participation.[89]

Key Messages Regarding Participants' Values and Preferences

Participants' values and preferences were conceptualized as willingness to be screened, motivators and barriers to participation in screening, and experienced based attitudes and perceptions of lung cancer screening. Participants reported a high willingness to be screened for lung cancer which may have been influenced by motivators such as smoking history, beliefs about early detection, familiarity with lung cancer and personal risk perception. Potential barriers to screening included perceptions of health care workers and settings, and travel and convenience of screening. Overall, participants' experiences with lung cancer screening were neutral or positive.

CQ4. *What is the optimal screening interval for screening for lung cancer?*

Four papers, which included three modelling studies[93-95] and one overview,[96] were found that addressed the question of optimal intervals for screening for lung cancer. Data incorporated into the models were drawn from the NLST trial,[93, 94] the UKLS study,[94] the PLCO study[93] and the Mayo Lung Project.[95] The overview paper[96] discussed NLST study data that were used to inform the American Association of Thoracic Surgery's guidelines on optimal screening intervals.

One study[93] applied five models (developed by investigators at five different institutions) calibrated to individual level data from the NLST and PLCO trials. The NLST study randomly assigned 26,722 participants to three annual LDCT screenings and 26,732 participants to three annual CXR screenings. The mixed gender sample ranged in age from 55 to 74 years old, and had a history of ≥30 pack-years of smoking. The PLCO study which included a mixed gender sample aged 55 to 64, randomly assigned 77,445 participants to annual CXR and 77,456 to usual care. In the models, three LDCT screening programs (triennial, biennial and annual screening) using a variety of start age, stop age, pack-years, and pack-years since quitting combinations were examined. Triennial screening produced minimal reductions in lung cancer mortality (range, 1.7% to 9.5% across models). Biennial screening programs generated lung cancer mortality reductions ranging from 2.3% to 14.8% across models. Annual screening produced considerably more benefits than other screening intervals, with a reduction in lung cancer mortality ranging from 4.3% to 39.1% across models. Additionally, the annual screening scenarios detected 48.1% to 56.9% of lung cancer cases at stage I/II. False positive findings increased proportionally to the additional CT scans required in each screening program, producing an average of 1.0 to 4.9 false positives per person screened. The authors assessed a screening approach comparable to that of the NLST (annual screening commencing at 55 years, ending at 80 years for ever-smokers with a history of ≥30 pack-years and ≤15 years since quitting) and determined it to be the most effective scenario with the ideal balance of benefits and harms.

Another modelling study[94] used the Liverpool Lung Project model to compare annual and biennial LDCT screening programs looking at two potential risk groups (the cohort recruited in the NLST study and the group selected in the UKLS pilot trial) and two outcomes of interest to this review (lung cancer mortality and overdiagnosis). The findings from the NLST study were used to predict the possible effects of a UK LDCT screening program. Differences in participant characteristics were minor; with the NLST trial including participants aged 55 to 74 years and the UKLS study including those aged 50 to 75 years. In terms of mortality, the results projected that annual LDCT screening with the UKLS eligibility criteria would prevent 956 (out of 330,000 screens, range 325 to 1,276) lung cancer deaths as opposed to biennial screening which would prevent 802 (out 180,000 screens, range 273 to 1,071) lung cancer deaths. Comparatively, annual screening with the NLST criteria would result in the prevention of 819 (range 278 to 1,093) lung cancer deaths while biennial screening would prevent 687 (range 234 to 917) lung cancer deaths. Considering harms, annual LDCT screening using UKLS criteria would result in 457 (range 0 to 748) overdiagnosed lung cancers from 330,000 screening episodes, while

biennial screening would yield 383 (range 0 to 627) overdiagnosed lung cancers from 180,000 screening episodes. Annual screening using the NLST eligibility criteria would produce 392 (range 0 to 641) overdiagnosed lung cancers while biennial screening would result in 329 overdiagnosed lung cancers (range 0 to 538). The authors concluded that estimates using annual LDCT screening with high risk populations appear to be the most effective in reducing lung cancer mortality and lowering rates of overdiagnosis. They also suggested additional research is required on screening intervals longer than one year.

The final modelling study[95] used the Markov Chain Monte Carlo approach and applied data from the Mayo Lung Project to estimate lead and sojourn time in lung cancer screening. The Mayo Lung Project included male heavy smokers aged 44 to 74 years (only data from aged 45 to 69 years were used in the model) who were given a screening test every four months. Screening tests included a combination of CXR and three-day pooled SC sampling. The authors used the model to estimate the results for male heavy smokers being screened at various intervals starting at age 45. The time between screening intervals was six, nine, 12 and 18 months. Results indicated that for a heavy smoker, a 12 month screening interval commencing at age 45 years and continuing until age 75 years would result in a 25.30% chance of lung cancer not being detected early if it is developed during that time period. The possibility of early detection not happening decreases to 16.79% when the screening interval is every six months. The authors concluded that increased frequency of screening will result in a longer average lead time which theoretically would lead to the lung cancer being treated at an earlier stage, presumably increasing the likelihood of survival.

The American Association of Thoracic Surgery's overview paper on lung screening guidelines[96] discussed the outcomes of the NLST study and echoed the conclusions drawn from the previously discussed modelling studies. The Association endorsed the use of annual LDCT screening, noting the 20% reduction in lung cancer deaths after three scans. The authors cautioned that baseline cancer risk does not decline after the initial three years, but actually steadily increases with each year of aging and/or exposure to smoke. As a result, they suggested that extending the annual screening phase from three to seven years warrants consideration. They posit that if active screening had been maintained throughout the four year observation period of the NLST trial, more early stage lung cancers would have been detected resulting in a greater reduction in lung cancer related deaths. The authors concluded that low risk of radiation exposure, minimal cost, and the reduced possibility of developing later stage lung cancer all support the use of annual LDCT screening.

Key Messages Regarding Screening Intervals

Overall, results from the abovementioned modelling studies indicate that annual LDCT screening, when compared to longer screening intervals, appears to be most effective, increasing diagnoses of lung cancer at earlier and more treatable stages thereby contributing to a greater reduction in lung cancer mortality. Additionally, annual screening programs appear to result in a lower rate of overdiagnosis.

CQ5. *What risk assessment tools are identified in the literature to assess the risk of lung cancer?*

Our search of the literature for evidence addressing the contextual questions did not locate any papers that presented lung cancer risk assessment tools.

CQ6. *What is the evidence that subgroups (Aboriginal populations, rural or remote populations, other ethnic populations) have a higher burden of disease, a differential treatment response, differential performance of screening tests, or barriers to implementation?*

Fourteen papers were found that addressed the question of subgroup differences.[97-109] Eleven papers addressed the issue of burden of lung cancer;[98-100, 102-109] eight of these addressed Canadian Aboriginal populations,[102-109] one looked at rural or remote populations,[104] and one considered burden in other ethnic populations.[104] Several other papers were included that examined burden associated with socio-economic status and sex.[98-100, 104] No evidence relating to differential performance of screening tests was found. One article included information on treatment response[97] and two reported on barriers to screening implementation.[101, 103]

Burden of Disease

Aboriginal Populations

According to a 2013 Canadian Partnership Against Cancer report on First Nations cancer control in Canada, lung cancer is the most common type of cancer diagnosed across Aboriginal peoples.[103] Despite this fact, there is little specific information on the burden of disease for lung cancer among this population. Our search of the peer-reviewed literature did not find any relevant studies from the last five years. The information available is largely from grey literature sources, primarily government reports, and is limited in scope. A major review of disparities in cancer control in Canada summarizes the issue: "Currently, there is no means to identify First Nations, Métis, and Inuit cancer patients in the cancer control system because this information is not recorded in cancer registries nor consistently in health care records."[104] As it is estimated that more than 85% of lung cancer cases are related to smoking,[109] smoking rates among Aboriginal Canadians are of particular concern. A 2013 report by Physicians for a Smoke-Free Canada, using data from the Canadian Community Health Survey from 2007-2010, found that the overall smoking rate for Aboriginal Canadians was 39% compared to 20.5% for non-Aboriginal Canadians.[108] Across Aboriginal groups, rates were highest among the Inuit (49%), followed by First Nations (40%) and Métis (37%). A more recent report on tobacco use in Canada that used data from Health Canada national surveys indicated that 16.1% of Canadians (approximately 4.6 million) were current smokers in 2012.

An Alberta Health Services report from 2012 on cancer screening in Alberta's Aboriginal population states that the three leading causes of cancer-related deaths are lung cancer, colon

cancer, and breast cancer, accounting for 19.5%, 9.2%, and 8.8% of cancer-related deaths among First Nations people, respectively.[105]

A brief report on lung and bronchus cancer by the Nunavut (84% Inuit population) government using data from 1999-2010, found that age-standardized rates for these cancers per 100,000 people have declined from a high of 320.96 in 2002 to 164.33 in 2010, but continue to significantly exceed the national rate of 55.00.[102] The average age of diagnosis was 66.16 years and 62% of those diagnosed died within one year of detection. Considering a major risk factor for lung cancer, 60% of the population of Nunavut aged ≥12 years self-reported as current smokers, which is more than three times the national average.

Finally, a Canadian Partnership Against Cancer report on a cancer control plan for Canadian First Nation, Métis and Inuit noted 70% of the adult Inuit population reported in 2006 that they smoked and that at the time of a circumpolar review of Inuit, Canadian Inuit had the highest lung cancer rates in the world. The data used in the original review[106] is dated (1989-2003) but these results are frequently cited in more recent reports.[107]

Rural and Remote

Our search of the peer-reviewed literature did not find any articles from the last five years dealing with differential burden of lung cancer in rural and remote versus urban areas in Canada. The information presented here is gathered from grey literature, primarily government reviews and reports. A major system performance review of disparities in cancer control in Canada that used 2007 data from provincial cancer agencies, Statistics Canada and the Canadian Cancer Registry, found that Canadians living in rural, remote or very remote areas had higher age-standardized incidence and mortality rates for lung cancer than those living in urban areas (incidence rate per 100,000 population: very remote 62.8, remote 58.6, rural 62.9, urban 55.0; mortality rates per 100,000 population: very remote 51.2. remote 48.1, rural 47.5, urban 43.1).[104] While the national review found no clear geographic pattern for early stage lung cancer, rates of advanced stage lung cancer increased with increasing remoteness of the location (incidence rates per 100,000 population: very remote 39.5, remote 35.3, rural 31.6, urban 31.9). The authors suggested this trend may, at least partially account for higher mortality rates in rural and remote locations. Considering a major risk factor for lung cancer, smoking rates are also higher in rural and remote areas. In 2011, the percentage of Canadians aged ≥12 years reporting daily or occasional smoking ranged from 24.0 in very remote areas to 19.3 in urban areas.

Other Ethnic Populations

Our search of the peer-reviewed literature did not find any articles from the last five years dealing with differential burden of lung cancer in other ethnic populations in Canada. The 2014 disparities in cancer control review did report on disparities by immigration status for some risk factors for lung cancer, but did not provide lung cancer specific incidence or mortality rates for any of these groups.[104]

Socio-economic Status and Sex

While not specifically identified in the contextual question, differences in the burden of lung cancer by socio-economic status (SES) and sex are briefly discussed as these subgroups were frequently mentioned in the available literature.[99, 100, 104]

Using data from the Ontario Cancer Registry, a study by Booth et al.[98] examined the impact of SES on stage of lung cancer at diagnosis and survival. The study did not find a significant association between SES and stage of disease at diagnosis for NSCLC. The comparison between the lowest SES quintile and the other four combined found 13% of both groups were diagnosed with stage I (P=0.406) and 47% versus 49% (P=0.204) were diagnosed with stage IV. The authors found substantial differences across income quintiles in five-year overall survival, for NSCLC (3% P=0.002), with higher income quintiles having the advantage; however, three-year cancer-specific survival difference was not significant (2% P=0.317). They concluded that SES remains associated with survival among cancer patients in Ontario but that it cannot be explained by differences in the stage of cancer at diagnosis. On the other hand, the Canadian Cancer Disparities Report[104] did identify some significant differences relating to lung cancer burden based on income quintile. Specifically this report indicated that people from low-income neighbourhoods have higher lung cancer incidence and mortality rates, higher rates of advanced-stage lung cancer at diagnosis and higher rates of smoking than do residents of higher-income neighbourhoods. In 2007 the age-standardized incidence rates of lung cancer per 100,000 population, by neighbourhood income quintile, from lowest to highest were 69.6, 61.0, 57.3, 52.3, and 43.1. The corresponding age-standardized mortality rates were 54.7, 48.2, 44.2, 40.6, 35.5.[104]

There are significant differences in lung cancer incidence rates between males and females in Canada that primarily reflect historical differences in tobacco use. According to the 2014 Canadian Cancer Statistics Report by the Canadian Cancer Society,[99] while men continue to have higher incidence rates than women (59 versus 48 per 100,000), the rate began leveling off in the mid-1980s with declines of about 2% per year. The incidence rate of lung cancer among Canadian women has not increased since 2006 and, given that tobacco consumption in this population began to decrease in the mid-1980s, incidence rates are expected to decline over the next 20 years. Overall the five-year relative survival ratio for lung cancer is poor (17%) compared to other cancers; however, there are significant sex-specific differences in this measure, with relative survival ratios of 20% for women and 14% for men. The reasons for this difference are unclear but it has been suggested that sex differences in the response to treatment (discussed briefly below) may contribute to variations in the mortality rate.[110]

Treatment Response

Only one article, a review from 2013, addressed the issue of differential treatment response and the focus was on lung cancer in women.[97] The authors point out that although treatment protocols for surgery, radiation therapy and cytotoxic chemotherapy are not sex-specific, women are more likely than men to receive molecularly targeted therapy for locally advanced and metastatic adenocarcinoma. They cited a study (the IDEAL 1 Trial, 2003) of pharmacological

treatment with the small molecule tyrosine kinase inhibitors (TKI) gefitinib and erlotinib which demonstrated improved response rates in women.[111] The IDEAL trial found that the odds of responding to this (2nd line) treatment was over 2.5 times higher for females than males (95% CI 1.19, 5.91; P=0.17).

Analysis of data from the American Surveillance Epidemiology and End Results registry of 228,572 lung cancer patients between 1975 and 1999 showed improved overall two and five-year survival rates in women relative to men.[97] In addition, women had significantly higher stage-specific survival. This advantage appeared to hold true for both small cell and non-small cell lung cancers. The authors suggested that these sex differences in outcomes indicate the possibility of sex-specific differences in tumors that may necessitate different treatment approaches.[97]

Barriers to Implementation

What is most notable about the literature on barriers to implementation of lung cancer screening is that, with the possible exception of the stigma attached to a history of smoking, the barriers for subpopulations of interest are similar to screening for other cancers or diseases.

<u>Aboriginal Populations, Rural and Remote, Other Ethnic Populations</u>

Our search of the peer-reviewed literature for the last five years did not locate any articles that specifically discussed lung cancer screening barriers in the subpopulations of interest. However, the topic of screening barriers in general is well covered in the grey literature. The First Nations Cancer Control Baseline Report[103] provides a substantive discussion of barriers to health service delivery including cancer screening generally. The barriers identified in the report include: access to rural and remote communities; coordination of care; patient identification; and community awareness and education (which includes cultural and language issues). Not surprisingly, barriers for rural and remote populations center largely on access to screening facilities and the availability of family doctors. Another source reported common barriers to screening among new immigrants including: lack of education and awareness about screening; lack of culturally appropriate screening services; and lack of English (or French) language proficiency.[101]

Key Messages Regarding Subgroup Differences

The majority of lung cancer cases are so closely linked to tobacco use that variations in burden of the disease in the identified subpopulations are largely a reflection of tobacco use rates among those groups. We found no recent evidence on differential performance of screening tests by subpopulation. Only one study discussed differential lung cancer treatment response and it suggested women may respond better than men to very specific molecular treatments. Barriers to implementation of lung cancer screening in the identified subpopulations are substantial but do not differ significantly from those identified for other types of cancer screening programs.

CQ7. *What is the cost-effectiveness of screening for lung cancer?*

Four studies were found that used modelling methodologies[112-114] or registry data[115] that addressed the question regarding cost-effectiveness of screening for lung cancer.

The most cited study conducted by a group of US-based researchers[112] used simulated cohorts (n=500,000), to assess the cost-effectiveness of three screening and/or preventive programs: computed tomography (CT) screening; smoking cessation; or CT and smoking cessation combined. The Lung Cancer Policy Model was used to predict the long-term effectiveness of screening, using cost of quality-adjusted life-year ($/QALY) gained as the outcome of interest. The findings indicate that for individuals with a history of 20 pack-years of smoking, the cost of annual screening beginning at 50 years of age was between US$126,000 and $169,000/QALY, and for individuals with a history of 40 pack-years the cost was between US$110,000 and $166,000/QALY. Of the three programs, the most cost-effective was annual CT screening plus smoking cessation therapy, with a cost of US$130,500 to $159,700/QALY when beginning at age 50 years, with a history of 20 pack-years.[112]

Two other US studies examined the cost-effectiveness of lung cancer screening.[113, 114] The first study used data from the NLST trial and performed an economic analysis using a budget impact model; a model that provided an estimate of screening costs per lung cancer death avoided. With an estimated 50% to 75% uptake of screening, the authors determined that annual LDCT screening for 10 years will add US$1.3 to $2.0 billion dollars to national health care expenditures.[113] Additionally, the authors noted that LDCT screening, at an uptake of 75% would avoid 8,100 premature deaths due to lung cancer annually. Finally, the cost of LDCT screening to avoid one lung cancer death is US$240,000; however, this would depend on factors such as prevalence of smokers who qualify for screening, screening uptake rates and costs of LDCT scans.[113] The second study examined the cost-utility of annual LDCT screening (with or without smoking cessation) for up to 15 years, using simulated models of a hypothetical cohort of 18 million high risk adults (30 pack-year smokers) aged 50 to 64 years.[114] The study concluded that, for this sample, LDCT screening alone would cost the health care system US$27.8 billion over 15 years, yielding 985,284 QALYs gained for a cost-utility ratio of US$28,240/QALY gained.[114] Adding smoking cessation therapies resulted in increased costs and QALYs saved, ranging from US$16,198 to $23,185/QALY. The authors concluded that screening with LDCT in addition to smoking cessation therapies is cost-effective.[114]

A study using registry data from the DLCST trial (n=4,104) assessed the health care costs and utilization of participants related to screening in the DLCST.[115] Trial participants were randomized to either five annual LDCT screening scans or usual care. The annual median cost per participant in the screening group was €1,342, compared to the usual care group at €1,190 (P<0.0001).[115] The authors note that despite the increase in costs related to screening, the expenditures were outweighed by the benefit of the true-negative group showing no significant increases in health care expenditures.[115]

Key Messages Regarding Cost-Effectiveness of Screening

It may be difficult to ascertain absolute estimates of cost (i.e., benefit, effectiveness) when considering screening for lung cancer, due to the heterogeneity across health care systems, outcomes assessed (QALYs, cost $/€), and assumptions made about characteristics of the

hypothetical cohorts (i.e., duration of heavy smoking, despite anti-smoking awareness campaigns and policies). However, Tota et al.[116] suggest that in order to improve the cost-effectiveness of screening for lung cancer with LDCT, the following conditions need to be considered: screening offers former smokers and current smokers with the greatest opportunity to reduce the risk of lung cancer; the uptake of screening activities needs to be optimized in high risk populations; and successful smoking cessation programs coupled with LDCT improve the cost-effectiveness of screening for lung cancer in high risk adults.[116]

Chapter 4: Discussion, Limitations and Conclusion

Discussion

To our knowledge this is the most up to date and comprehensive review on both the benefits and harms of lung cancer screening. To address the key questions about the benefits and harms of lung cancer screening in the population of interest (adults aged ≥18 years who are at average or high risk but are not suspected of having lung cancer), a reasonable amount of direct, high level (mostly RCT) and low GRADE quality evidence was found. Thirty-three studies conducted in the US and Europe over the last 50 years, that used multiple strategies for lung cancer screening (CXR, CXR plus SC, LDCT) in general adult populations (diversity in terms of age, smoking history and gender) comprised the body of evidence available for this review.

Benefits of Screening

For the critical outcomes of lung cancer mortality and all-cause mortality the available evidence indicated there is no benefit of CXR screening, with or without SC, when compared to no screening or less intensive screening. Pooled analyses of preliminary results from three relatively small trials comparing LDCT to usual care in high risk adults found no significant benefits for mortality with five years or less follow-up. One high quality trial with a large sample of high risk adults (NLST) and a median follow-up of 6.5 years found screening with LDCT showed significant benefits for mortality when compared with screening with CXR [lung cancer mortality RR 0.80 (95% CI 0.70, 0.92), NNS 308 (95% CI 201, 787); all-cause mortality RR 0.94 (95% CI 0.88, 1.00), NNS 219 (95% CI 115, 5,556)].[10] Of the two studies that examined subgroup populations (i.e., former smokers, never smokers, current smokers, men, women, age cohorts), no statistically significant differences were observed in terms of lung cancer mortality between the (CXR or CXR plus SC) screened and unscreened participants.

Consistent with expectations around the efficacy of screening, most screening strategies for lung cancer showed significant benefits in terms of disease detection. For CXR and LDCT screening, more cases of early stage NSCLC and fewer cases of late stage malignancy were observed in the screened and more intensively screened groups compared to the control groups; with the exception of early stage disease detection using dual testing with CXR and SC compared to less intensive screening. LDCT demonstrated better efficacy than CXR in a sample of high risk adults, detecting significantly more cases of early stage disease [LDCT 57.0% versus CXR 39.1%; RR 1.46 (95% CI 1.33, 1.61)] and significantly fewer cases of late stage disease [LDCT 43.0% versus CXR 60.9%; RR 0.71 (95% CI 0.65, 0.77)].[10]

We found little evidence for the outcome of smoking cessation rate. A similar pattern was observed across the available evidence; none of the studies found a difference in smoking cessation rates between the screened group and the control group after various time points (one, two, five, six and 28 years).

Little and inconsistent evidence was found regarding incidental findings of lung cancer screening.

Harms of Screening

The evidence for harms was primarily obtained from observational studies resulting in low GRADE quality of evidence.

For CXR, of all cases of lung cancer diagnosed in the screened population an estimated 2.27% to 16.28% were overdiagnosed. Complications of invasive follow-up procedures resulted in 28.60 deaths (95% CI 16.02, 41.17) per 1,000 patients who underwent invasive follow-up testing as a result of positive screening tests; an even higher proportion of deaths were observed when SC was added to CXR as part of the screening protocol [47.67 deaths (95% CI 23.86, 71.49) per 1,000]. Moreover, for every 1,000 patients who underwent invasive follow-up procedures, 63.32 patients (95% CI 42.92, 92.49) experienced major complications requiring hospitalization/medical intervention (e.g., hemothorax and pneumothorax requiring tube placement, lung collapse, severe pain, cardiac arrhythmias and thromboembolic complications[73, 74]). The median absolute number of false positives observed for CXR testing was 65.0 (range 34.0 to 136.7) per 1,000 adults screened, and as part of diagnostic follow-up, individuals with benign conditions were subjected to both minor and major invasive procedures at a rate of 2.30 (95% CI 1.49, 3.11) and 2.73 (95% CI 0.96, 4.51) individuals per 1,000 screened, respectively.

LDCT screening was also associated with both critical and important harms. Overdiagnosis ranged from 10.99% to 25.83%. Invasive follow-up procedures performed as a result of positive screening tests were associated with 11.18 deaths (95% CI 5.07, 17.28) and 43.29 patients experiencing major complications (95% CI 32.00, 54.58) per 1,000 patients who underwent invasive follow-up testing. The median absolute number of false positives associated with a baseline or single LDCT screen was 167.1 (range 79.0 to 255.3) per 1,000 participants tested and 233.0 (range 6.4 to 690.0) per 1,000 screened with multiple rounds of screening. As part of diagnostic follow-up, individuals with benign conditions were subjected to both minor and major invasive procedures at a rate of 7.16 (95% CI 3.27, 11.05) and 4.98 (95% CI 3.68, 6.29) individuals per 1,000 screened, respectively.

There was little and inconsistent evidence addressing the other harms of interest to this review (anxiety, quality of life, infection and bleeding from invasive follow-up testing).

Test Properties and Performance

For test properties and performance we only looked at LDCT, the screening strategy that showed benefit for mortality outcomes. Test properties varied across studies depending on the type of reference standard applied, cut-off or threshold value for a positive test, and LDCT technology and technique used. Sensitivity of LDCT ranged from 80% to 100% and specificity ranged from 28% to 100%. None of the included studies directly compared various LDCT technologies or varying protocols used by radiologists and their effect on test performance. In the available evidence, diagnostic test properties were highest overall with the multi slice detector along with computer assisted reading/diagnosis plus two independent radiologist readers.

Contextual Evidence

Surveyed adults reported a high willingness to be screened for lung cancer which may have been influenced by motivators such as smoking history, beliefs about early detection, familiarity with lung cancer and personal risk perception. Potential barriers to screening included perceptions of health care workers and settings, and travel and convenience of screening. Overall, participants' experiences with lung cancer screening were neutral or positive.

The majority of lung cancer cases are so closely linked to tobacco use that variations in burden of the disease in Canadian rural, remote, Aboriginal and other ethnic populations are largely a reflection of tobacco use rates among those groups. We found no recent evidence on differential performance of screening tests by subpopulation. Only one study discussed differential lung cancer treatment response and it suggested women may respond better than men to very specific molecular treatments. Barriers to implementation of lung cancer screening in rural, remote, Aboriginal, and other ethnic populations are substantial but do not differ significantly from those identified for other types of cancer screening programs.

Results from modelling studies indicate that annual LDCT screening, when compared to longer screening intervals, appears to be most effective, increasing diagnoses of lung cancer at earlier and more treatable stages thereby contributing to a greater reduction in lung cancer mortality. Additionally, annual screening programs appear to result in a lower rate of overdiagnosis.

Estimating the absolute costs of lung cancer screening is complicated by heterogeneity across health care systems, variety in outcomes (QALYs, cost \$/€), and different assumptions about hypothetical cohorts (e.g., duration of heavy smoking, despite anti-smoking awareness campaigns and policies). Modelling studies suggest that annual LDCT screening, when compared to longer screening intervals, appears to be most cost-effective.

Limitations

There was substantial variability observed across the included studies in terms of samples, screening tests, outcomes, comparators, length of follow-up, locations and timing. Of the 13 RCTs contributing to the benefits of screening, six trials included men only. While most studies targeted current and former heavy smokers, three studies included some never smokers; it is important to acknowledge that socially undesirable behaviours, including smoking, are notoriously under reported. The enrollment age spanned over 40 years (range 35 to 74 years) with about half of the studies including some participants aged <50 years and about three-quarters including some participants aged >70. The particular screening test or tests differed across studies (CXR, SC and/or LDCT) and control groups also varied (no screening, usual care, less intensive protocols using the same test(s)). The screening interventions varied based on available technology, most likely indicative of each country's health care system and access to screening expertise and equipment at the time of each study (a period which spans the last half century). Moreover, the countries where and the decades when these studies took place varied greatly in regards to the social acceptability of smoking and the existence of public health

policies, campaigns, and laws related to smoking. Finally, for mortality outcomes, the generalizability of the results of LDCT screening trials is limited by considerable variation in and relatively short length of follow-up (three studies had follow-up of five years or less).

There were a number of gaps and paucity within the available evidence. This review found only a few studies that reported on some of the important outcomes (e.g., smoking cessation, incidental findings, anxiety and quality of life). Very few papers included sub-analyses to fully address the question about sub-group differences that may influence the underlying risk of lung cancer. Most of the harms data was obtained from observational studies. Within the GRADE assessments, publication bias could not be evaluated, given the low number of included studies. The section on test properties was limited to LDCT screening. Our search for recent contextual evidence located no lung cancer risk assessment tools. Finally, we restricted our searches to papers in English or French, thus we may have missed the opportunity to analyze data from papers written in other languages.

Conclusion

This updated review with comprehensive evidence draws similar conclusions to those of the most recent USPSTF and Cochrane reviews on lung cancer screening.[25, 26] Considering lung cancer and all-cause mortality outcomes, the available evidence does not support lung cancer screening with CXR, with or without SC, in average to high risk adults.

Screening for lung cancer with LDCT showed potential benefit in a large sample of high risk adults when compared with CXR and reduced lung cancer mortality by 20% and all-cause mortality by 6%. Future long-term follow-up data from on-going LDCT trials[20, 21, 37, 40, 49, 117] will provide more conclusive evidence on the effectiveness of LDCT screening and further insights into optimal age for screening, screening interval and frequency. Finally, it is important to acknowledge the test's poor specificity and associated harms, including overdiagnosis and false positives as well as major complications and death resulting from invasive follow-up procedures, pose challenges for clinicians and public health professionals to implement screening and warrant the need to develop standardized practices.

References

1. Palda VA, Van Spall HGC, Canadian Task Force for Preventive Health Care, Canadian Task Force for Preventive Health Care. Screening for lung cancer: Updated recommendations from the Canadian Task Force on Preventive Health Care. Ottawa, ON; 2003. Available from: http://canadiantaskforce.ca/perch/resources/update-recommendations.pdf.

2. American Cancer Society (ACS) [Internet]. Lung cancer. Atlanta, GA: American Cancer Society, Inc.; 2014. Available from: http://www.cancer.org/cancer/lungcancer/index.

3. The Lung Association [Internet]. Lung cancer: what is lung cancer? Ottawa, ON: Canadian Lung Association; 2012. Available from: http://www.lung.ca/diseases-maladies/cancer-cancer/what-quoi/index_e.php.

4. Canadian Cancer Society's Advisory Committee. Canadian cancer statistics 2013. Toronto, ON; 2013. ISSN 0835-2976. Available from: http://www.cancer.ca/en/cancer-information/cancer-101/canadian-cancer-statistics-publication/?region=on.

5. Crawford, S M, Sauerzapf, V, Haynes, R, et al. Social and geographical factors affecting access to treatment of colorectal cancer: A cancer registry study. BMJ Open. 2012; 2(2):e000410. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3341592/.

6. Canadian Cancer Society [Internet]. Risk factors for lung cancer. Toronto, ON: Canadian Cancer Society; 2014. Available from: http://www.cancer.ca/en/cancer-information/cancer-type/lung/risks/?region=on.

7. Health Canada [Internet]. Canadian Tobaco Use Monitoring Survey (CTUMS) 2012. Ottawa, ON: Health Canada; 2013. Available from: http://www.hc-sc.gc.ca/hc-ps/tobac-tabac/research-recherche/stat/ctums-esutc_2012-eng.php.

8. Lung cancer: what are the risk factors? Atlanta, GA: Centers for Disease Control and Prevention. Available from: http://www.cdc.gov/cancer/lung/basic_info/risk_factors.htm.

9. U. S. Preventive Services Task Force. Lung cancer screening: recommendation statement. Ann Intern Med. 2004; 140(9):738-9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/?term=15126258.

10. National Lung Screening Trial Research Team, Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med. 2011; 365(5):395-409. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21714641.

11. Moyer VA, U. S. Preventive Services Task Force. Screening for lung cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2014; 160(5):330-8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/?term=24378917.

12. Wender R, Fontham ETH, Barrera E, Jr., Colditz GA, Church TR, Ettinger DS, et al. American Cancer Society lung cancer screening guidelines. CA Cancer J Clin. 2013; 63(2):107-17. Available from: http://www.ncbi.nlm.nih.gov/pubmed/?term=23315954.

13.   Alberts WM, American College of Chest Physicians. Introduction: Diagnosis and management of lung cancer: ACCP evidence-based clinical practice guidelines (2nd edition). Chest. 2007; 132(Suppl 3):S20-2. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17873157.

14.   American Lung Association [Internet]. Providing guidance on lung cancer screening to patients and physicians. Washington, DC; 2013. Available from: http://www.lung.org/lung-disease/lung-cancer/lung-cancer-screening-guidelines/lung-cancer-screening.pdf.

15.   Jaklitsch MT, Jacobson FL, Austin JHM, Field JK, Jett JR, Keshavjee S, et al. The American Association for Thoracic Surgery guidelines for lung cancer screening using low-dose computed tomography scans for lung cancer survivors and other high-risk groups. The Journal of Thoracic and Cardiovascular Surgery. 2012; 144(1):33-8. Available from: http://www.sciencedirect.com/science/article/pii/S0022522312006009.

16.   National Comprehensive Cancer Network [Internet]. NCCN guidelines for patients. Fort Washington, PA; 2014. Available from: http://www.nccn.org/patients/guidelines/lung_screening/index.html.

17.   Roberts H, Walker-Dilks C, Sivjee K, Ung Y, Yasufuku K, Hey A, et al. Screening high-risk populations for lung cancer. Toronto, ON; 2013. Program in Evidence-based Care Evidence-based Series No.: 15-10. Available from: https://www.cancercare.on.ca/common/pages/UserFile.aspx?fileId=287879.

18.   van Iersel CA, de Koning HJ, Draisma G, Mali WPTM, Scholten ET, Nackaerts K, et al. Risk-based selection from the general population in a screening trial: Selection criteria, recruitment and power for the Dutch-Belgian randomised lung cancer multi-slice CT screening trial (NELSON). Int J Cancer. 2007; 120(4):868-74. Available from: http://dx.doi.org/10.1002/ijc.22134.

19.   Saghir Z, Dirksen A, Ashraf H, Bach KS, Brodersen J, Clementsen PF, 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):296-301. Available from: http://www.ncbi.nlm.nih.gov/pubmed/?term=22286927.

20.   Lopes Pegna A, Picozzi G, Mascalchi M, Maria Carozzi F, Carrozzi L, Comin C, 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. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18723240.

21.   Infante M, Cavuto S, Lutman FR, Brambilla G, Chiesa G, Ceresoli G, et al. A randomized study of lung cancer screening with spiral computed tomography: three-year results from the DANTE trial. Am J Respir Crit Care Med. 2009; 180(5):445-53. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19520905.

22.   Infante M, Lutman FR, Cavuto S, Brambilla G, Chiesa G, Passera E, et al. Lung cancer screening with spiral CT: Baseline results of the randomized DANTE trial. Lung Cancer. 2008; 59(3):355-63. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17936405.

23.   Pastorino U, Rossi M, Rosato V, Marchiano A, Sverzellati N, Morosi C, 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):308-15. Available
from: http://www.ncbi.nlm.nih.gov/pubmed/?term=22465911.

24.   Pastorino U, Bellomi M, Landoni C, De Fiori E, Arnaldi P, Picchio M, et al. Early lung-
cancer detection with spiral CT and positron emission tomography in heavy smokers: 2-year
results. The Lancet. 2003; 362(9384):593-7. Available
from: http://www.ncbi.nlm.nih.gov/pubmed/12944057.

25.   Manser R, Lethaby A, Irving LB, Stone C, Byrnes G, Abramson MJ, et al. Screening for lung
cancer. Cochrane Database Syst Rev. 2013; 6:CD001991. Available
from: http://www.ncbi.nlm.nih.gov/pubmed/23794187.

26.   Humphrey L, Deffebach M, Pappas M, Baumann C, Artis K, Priest Mitchell J, et al.
Screening for lung cancer: systematic review to update the U.S. Preventive Services Task
Force Recommendation. Rockville, MD; 2013. 13-05188-EF-1. Available
from: http://www.ncbi.nlm.nih.gov/books/NBK154610/.

27.   Distiller (DistillerSR systematic review software) [computer program] [program]. Ottawa,
ON: Evidence Partners. Available from: http://systematic-review.net/.

28.   Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, et al. GRADE guidelines: 2.
Framing the question and deciding on important outcomes. J Clin Epidemiol. 2011;
64(4):395-400. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21194891.

29.   Higgins JPT, Altman DG, Sterne JAC, Group TCSM, Group TCBM. Assessing risk of bias
in included studies. In: J.P.T. Higgins, S. Green, editors. *Cochrane Handbook for Systematic
Reviews of Interventions Version 5.1.0 [updated March 2011]*: The Cochrane Collaboration;
2011.

30.   GRADEpro. [Computer program]. Version 3.2 for Windows [program]. Available
from: http://ims.cochrane.org/revman/other-resources/gradepro/download.

31.   GRADE working group. Place published unknown: GRADE working group; 2005. Available
from: http://www.gradeworkinggroup.org/.

32.   Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE:
An emerging consensus on rating quality of evidence and strength of recommendations. BMJ
(Clinical research ed). 2008; 336(7650):924-6. Available
from: http://www.ncbi.nlm.nih.gov/pubmed/?term=18436948.

33.   DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986; 7(3):177-
88. Available from: http://www.ncbi.nlm.nih.gov/pubmed/?term=3802833.

34.   Wallis S. Binomial confidence intervals and contingency tests: mathematical fundamentals
and the evaluation of alternative methods. JQL. 2013; 20(3):178-208. Available
from: http://corplingstats.wordpress.com/2012/03/31/binomial-distributions/.

35.   Review Manager (RevMan) [Computer Program]. Version 5.1 [program]. Copenhagen: The
Nordic Cochrane Centre, The Cochrane Collaboration. Available
from: http://tech.cochrane.org/revman/download.

36.   Stata statistical software [Computer Program]: release 12 [program]. Station, TX: StataCorp LP. Available from: http://www.stata.com/.

37.   Horeweg N, van der Aalst CM, Thunnissen E, Nackaerts K, Weenink C, Groen HJM, et al. Characteristics of lung cancers detected by computer tomography screening in the randomized NELSON trial. Am J Respir Crit Care Med. 2013; 187(8):848-54. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23348977.

38.   Oken MM, Hocking WG, Kvale PA, Andriole GL, Buys SS, Church TR, et al. Screening by chest radiograph and lung cancer mortality: the Prostate, Lung, Colorectal, and Ovarian (PLCO) randomized trial. JAMA. 2011; 306(17):1865-73. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22031728.

39.   Marcus PM, Bergstralh EJ, Zweig MH, Harris A, Offord KP, Fontana RS. Extended lung cancer incidence follow-up in the Mayo Lung Project and overdiagnosis. J Natl Cancer Inst. 2006; 98(11):748-56. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16757699.

40.   Ashraf H, Saghir Z, Dirksen A, Pedersen JH, Thomsen LH, Døssing M, 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. Available from: http://thorax.bmj.com/content/early/2014/01/17/thoraxjnl-2013-203849.abstract.

41.   Gohagan J, Marcus P, Fagerstrom R, Pinsky P, Kramer B, Prorok P, et al. Baseline findings of a randomized feasibility trial of lung cancer screening with spiral CT scan vs chest radiograph: the Lung Screening Study of the National Cancer Institute. Chest. 2004; 126(1):114-21. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15249451.

42.   Kubik A, Haerting J. Survival and mortality in a randomized study of lung cancer detection. Neoplasma. 1990; 37(4):467-75. Available from: http://www.ncbi.nlm.nih.gov/pubmed/2234207.

43.   Baker RR, Tockman MS, Marsh BR, Stitik FP, Ball WC, Jr., Eggleston JC, et al. Screening for bronchogenic carcinoma: the surgical experience. The Journal of Thoracic and Cardiovascular Surgery. 1979; 78(6):876-82. Available from: http://www.ncbi.nlm.nih.gov/pubmed/502570.

44.   Dales LG, Friedman GD, Collen MF. Evaluating periodic multiphasic health checkups: a controlled trial. J Chronic Dis. 1979; 32(5):385-404. Available from: http://www.ncbi.nlm.nih.gov/pubmed/109452.

45.   Flehinger BJ, Kimmel M. The natural history of lung cancer in a periodically screened population. Biometrics. 1987; 43(1):127-44. Available from: http://www.ncbi.nlm.nih.gov/pubmed/3567302.

46.   Anon. Lung cancer detection by chest x-rays at 6 monthly intervals. N S Med Bull. 1970; 49(1):14-5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/?term=5262952.

47.   Mazzone PJ, Obuchowski N, Fu AZ, Phillips M, Meziane M. Quality of life and healthcare use in a randomized controlled lung cancer screening study. Annals of the American Thoracic Society. 2013; 10(4):324-9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/?term=23952850.

48. Lopes Pegna A, Picozzi G, Falaschi F, Carrozzi L, Falchini M, Carozzi FM, et al. Four-year results of low-dose CT screening and nodule management in the ITALUNG trial. Journal of Thoracic Oncology. 2013; 8(7):866-75. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23612465.

49. Becker N, Motsch E, Gross ML, Eigentopf A, Heussel CP, Dienemann H, et al. Randomized study on early detection of lung cancer with MSCT in Germany: study design and results of the first screening round. J Cancer Res Clin Oncol. 2012; 138(9):1475-86. Available from: http://www.ncbi.nlm.nih.gov/pubmed/?term=22526165.

50. Dominioni L, Rotolo N, Mantovani W, Poli A, Pisani S, Conti V, et al. A population-based cohort study of chest x-ray screening in smokers: lung cancer detection findings and follow-up. BMC Cancer. 2012; 12:18. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22251777.

51. Croswell JM, Baker SG, Marcus PM, Clapp JD, Kramer BS. Cumulative incidence of false-positive test results in lung cancer screening: a randomized trial. Erratum in: Ann Intern Med. 2010 Jun 1;152(11):759. Ann Intern Med. 2010; 152(8):505-12. Available from: http://www.ncbi.nlm.nih.gov/pubmed/?term=20404381.

52. Sone S, Nakayama T, Honda T, Tsushima K, Li F, Haniuda M, et al. Long-term follow-up study of a population-based 1996-1998 mass screening programme for lung cancer using mobile low-dose spiral computed tomography. Lung Cancer. 2007; 58(3):329-41. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17675180.

53. Crestanello JA, Allen MS, Jett JR, Cassivi SD, Nichols FC, 3rd, Swensen SJ, et al. Thoracic surgical operations in patients enrolled in a computed tomographic screening trial. The Journal of Thoracic and Cardiovascular Surgery. 2004; 128(2):254-9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15282462.

54. Rzyman W, Jelitto-Gorska M, Dziedzic R, Biadacz I, Ksiazek J, Chwirot P, et al. Diagnostic work-up and surgery in participants of the Gdansk lung cancer screening programme: the incidence of surgery for non-malignant conditions. Interact Cardiovasc Thorac Surg. 2013; 17(6):969-73. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24008181.

55. Byrne MM, Weissfeld J, Roberts MS. Anxiety, fear of cancer, and perceived risk of cancer following lung cancer screening. Med Decis Making. 2008; 28(6):917-25. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18725404.

56. Menezes RJ, Roberts HC, Paul NS, McGregor M, Chung TB, Patsios D, et al. Lung cancer screening using low-dose computed tomography in at-risk individuals: the Toronto experience. Lung Cancer. 2010; 67(2):177-83. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19427055.

57. Veronesi G, Bellomi M, Mulshine JL, Pelosi G, Scanagatta P, Paganelli G, et al. Lung cancer screening with low-dose computed tomography: a non-invasive diagnostic protocol for baseline lung nodules. Lung Cancer. 2008; 61(3):340-9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18308420.

58. MacRedmond R, McVey G, Lee M, Costello RW, Kenny D, Foley C, et al. Screening for lung cancer using low dose CT scanning: results of 2 year follow up. Thorax. 2006; 61(1):54-6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16396954.

59. Blanchon T, Brechot JM, Grenier PA, Ferretti GR, Lemarie E, Milleron B, 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):50-8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17624475.

60. Callol L, Roig F, Cuevas A, de Granda JI, Villegas F, Jareño J, et al. Low-dose CT: A useful and accessible tool for the early diagnosis of lung cancer in selected populations. Lung Cancer. 2007; 56(2):217-21. Available from: http://www.sciencedirect.com/science/article/pii/S0169500207000025.

61. Sobue T, Moriyama N, Kaneko M, Kusumoto M, Kobayashi T, Tsuchiya R, et al. Screening for lung cancer with low-dose helical computed tomography: anti-lung cancer association project. J Clin Oncol. 2002; 20(4):911-20. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11844811.

62. Henschke CI, McCauley DI, Yankelevitz DF, Naidich DP, McGuinness G, Miettinen OS, et al. Early Lung Cancer Action Project: a summary of the findings on baseline screening. The Oncologist. 2001; 6(2):147-52. Available from: http://theoncologist.alphamedpress.org/content/6/2/147.abstract.

63. Diederich S, Thomas M, Semik M, Lenzen H, Roos N, Weber A, et al. Screening for early lung cancer with low-dose spiral computed tomography: results of annual follow-up examinations in asymptomatic smokers. Eur Radiol. 2004; 14(4):691-702. Available from: http://www.ncbi.nlm.nih.gov/pubmed/14727146.

64. Wilde J. A 10 year follow-up of semi-annual screening for early detection of lung cancer in the Erfurt County, GDR. The European Respiratory Journal. 1989; 2(7):656-62. Available from: http://www.ncbi.nlm.nih.gov/pubmed/2776873.

65. Duffy SW, Field JK, Allgood PC, Seigneurin A. Translation of research results to simple estimates of the likely effect of a lung cancer screening programme in the United Kingdom. Br J Cancer. 2014; 110(7):1834-40. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24525696.

66. Yankelevitz DF, Kostis WJ, Henschke CI, Heelan RT, Libby DM, Pasmantier MW, et al. Overdiagnosis in chest radiographic screening for lung carcinoma: frequency. Cancer. 2003; 97(5):1271-5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/?term=12599235.

67. Patz EF, Jr., Pinsky P, Gatsonis C, Sicks JD, Kramer BS, Tammemagi MC, et al. Overdiagnosis in low-dose computed tomography screening for lung cancer. JAMA Intern Med. 2014; 174(2):269-74. Available from: http://www.ncbi.nlm.nih.gov/pubmed/?term=24322569.

68. Veronesi G, Maisonneuve P, Bellomi M, Rampinelli C, Durli I, Bertolotti R, et al. Estimating overdiagnosis in low-dose computed tomography screening for lung cancer: a cohort study. Ann Intern Med. 2012; 157(11):776-84. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23208167.

69.   Melamed MR, Flehinger BJ, Zaman MB, Heelan RT, Perchick WA, Martini N. Screening for Early Lung-Cancer - Results of the Memorial-Sloan-Kettering Study in New-York. Chest. 1984; 86(1):44-53. Available from: http://www.ncbi.nlm.nih.gov/pubmed/6734291.

70.   Fontana RS, Sanderson DR, Woolner LB, Taylor WF, Miller WE, Muhm JR. Lung cancer screening: the Mayo program. J Occup Med. 1986; 28(8):746-50. Available from: http://www.ncbi.nlm.nih.gov/pubmed/3528436.

71.   Kubik A, Parkin DM, Khlat M, Erban J, Polak J, Adamec M. Lack of benefit from semi-annual screening for cancer of the lung: follow-up report of a randomized controlled trial on a population of high-risk males in Czechoslovakia. Int J Cancer. 1990; 45(1):26-33. Available from: http://www.ncbi.nlm.nih.gov/pubmed/2404878.

72.   Petersen RH, Hansen HJ, Dirksen A, Pedersen JH. Lung cancer screening and video-assisted thoracic surgery. Journal of Thoracic Oncology. 2012; 7(6):1026-31. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22588154.

73.   Goulart BHL, Ramsey SD. Moving beyong the National Lung Screening Trial: Discussing strategies for implementation of lung cancer screening programs. The Oncologist. 2013; 18:941-6. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3755932.

74.   NLST. NLST medical complications of diagnostic evaluation for lung cancer: data dictionary. 2014. Available from: https://biometry.nci.nih.gov/cdas/attach/serve/datasetdocumentation/nlst/comprehensive/medical_complications/medical_complications.dictionary.d041614.pdf.

75.   Horeweg N, van der Aalst CM, Vliegenthart R, Zhao YR, Xie X, Scholten ET, et al. Volumetric computed tomography screening for lung cancer: three rounds of the NELSON trial. The European Respiratory Journal. 2013; 42(6):1659-67. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23845716.

76.   Swensen SJ, Jett JR, Hartman TE, Midthun DE, Mandrekar SJ, Hillman SL, et al. CT screening for lung cancer: five-year prospective experience. Radiology. 2005; 235(1):259-65. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15695622.

77.   Hocking WG, Tammemagi MC, Commins J, Oken MM, Kvale PA, Hu P, et al. Diagnostic evaluation following a positive lung screening chest radiograph in the Prostate, Lung, Colorectal, Ovarian (PLCO) Cancer Screening Trial. Lung Cancer. 2013; 82(2):238-44. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23993734.

78.   Wilson DO, Weissfeld JL, Fuhrman CR, Fisher SN, Balogh P, Landreneau RJ, et al. The Pittsburgh Lung Screening Study (PLuSS): outcomes within 3 years of a first computed tomography scan. Am J Respir Crit Care Med. 2008; 178(9):956-61. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18635890.

79.   van den Bergh KA, Essink-Bot ML, Borsboom GJ, Th Scholten E, Prokop M, de Koning HJ, et al. Short-term health-related quality of life consequences in a lung cancer CT screening trial (NELSON). Br J Cancer. 2010; 102(1):27-34. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19935789.

80. Taylor KL, Shelby R, Gelmann E, McGuire C. Quality of life and trial adherence among participants in the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial. J Natl Cancer Inst. 2004; 96(14):1083-94. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15265970.

81. Lacasse Y. Screening with chest radiography did not reduce lung cancer mortality in older patients. Ann Intern Med. 2012; 156(6):JC3-8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22431694.

82. Aberle DR, DeMello S, Berg CD, Black WC, Brewer B, Church TR, et al. Results of the two incidence screenings in the National Lung Screening Trial. N Engl J Medicine. 2013; 369(10):920-31. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24004119.

83. van Klaveren RJ, Oudkerk M, Prokop M, Scholten ET, Nackaerts K, Vernhout R, et al. Management of lung nodules detected by volume CT scanning. N Engl J Medicine. 2009; 361(23):2221-9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19955524.

84. Toyoda Y, Nakayama T, Kusunoki Y, Iso H, Suzuki T. Sensitivity and specificity of lung cancer screening using chest low-dose computed tomography. Br J Cancer. 2008; 98(10):1602-7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18475292.

85. Tsushima K, Sone S, Hanaoka T, Kubo K. Radiological diagnosis of small pulmonary nodules detected on low-dose screening computed tomography. Respirology. 2008; 13(6):817-24. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18811880.

86. Pallin M, Walsh S, O'Driscoll MF, Murray C, Cahalane A, Brown L, et al. Overwhelming support among urban Irish COPD patients for lung cancer screening by low-dose CT scan. Lung. 2012; 190(6):621-8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23064487.

87. Jonnalagadda S, Bergamo C, Lin JJ, Lurslurchachai L, Diefenbach M, Smith C, et al. Beliefs and attitudes about lung cancer screening among smokers. Lung Cancer. 2012; 77(3):526-31. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22681870.

88. Patel D, Akporobaro A, Chinyanganya N, Hackshaw A, Seale C, Spiro SG, et al. Attitudes to participation in a lung cancer screening trial: a qualitative study. Thorax. 2012; 67(5):418-25. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22106018.

89. van den Bergh KAM, Essink-Bot ML, van Klaveren RJ, de Koning HJ. Informed participation in a randomised controlled trial of computed tomography screening for lung cancer. The European Respiratory Journal. 2009; 34(3):711-20. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19282345.

90. Park ER, Streck JM, Gareen IF, Ostroff JS, Hyland KA, Rigotti NA, et al. A qualitative study of lung cancer risk perceptions and smoking beliefs among National Lung Screening Trial participants. Nicotine and Tobacco Research. 2014; 16(2):166-73. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23999653.

91. Tanner NT, Egede LE, Shamblin C, Gebregziabher M, Silvestri GA. Attitudes and beliefs toward lung cancer screening among US veterans. Chest. 2013; 144(6):1783-7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23764896.

92. Flynn AE, Peters MJ, Morgan LC. Attitudes towards lung cancer screening in an Australian high-risk population. Lung Cancer Intern. 2013; 2013: Article ID 789057, 7 pages. Available from: http://www.hindawi.com/journals/lci/2013/789057/.

93. de Koning HJ, Meza R, Plevritis SK, ten Haaf K, Munshi VN, Jeon J, et al. Benefits and harms of computed tomography lung cancer screening strategies: a comparative modeling study for the U.S. Preventive Services Task Force. Ann Intern Med. 2014; 160(5):311-20. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24379002.

94. Duffy SW, Field JK, Allgood PC, Seigneurin A. Translation of research results to simple estimates of the likely effect of a lung cancer screening programme in the United Kingdom. Br J Cancer. 2014; 110(7):1834-40. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24525696.

95. Wu D, Erwin D, Rosner GL. Sojourn time and lead time projection in lung cancer screening. Lung Cancer. 2011; 72(3):322-6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21075475.

96. Jaklitsch MT, Jacobson FL. Age limits and frequency of negative scans: When should lung cancer screening end? J Surg Oncol. 2013; 108(5):301-3. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24006235.

97. Graham PD, Thigpen SC, Geraci SA. Lung cancer in women. South Med J. 2013; 106(10):582-7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24096953.

98. Booth CM, Li G, Zhang-Salomons J, Mackillop WJ. The impact of socioeconomic status on stage of cancer at diagnosis and survival: A population-based study in Ontario, Canada. Cancer. 2010; 116(17):4160-7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20681012.

99. Canadian Cancer Society's Advisory Committee on Cancer Statistics. Canadian cancer statistics 2014. Toronto, ON; 2014. Available from: http://www.cancer.ca/en/cancer-information/cancer-101/canadian-cancer-statistics-publication/?region=on.

100. The Canadian Cancer Society of Saskatchewan [Internet]. Lung cancer: the facts. Saskatchewan, SK: Canadian Cancer Society. Available from: http://ccssk.convio.net/site/DocServer/Lung_cancer_facts_2013_EN_FINAL.pdf;jsessionid=E8DC79A420B303DF3723946330C0770D.app331a?docID=1921&autologin=true.

101. Canadian Partnership Against Cancer. Screening portfolio. Lung cancer screening. Expert Panel: summary of existing and new evidence. Toronto, ON; 2011. Available from: http://www.lungcancercanada.ca/resources/site1/general/PDF/CPAC_Lung_Cancer_Screening_FINAL.pdf.

102. Nunavut [Internet]. Nunavut 1999-2010: lung and bronchus cancer. Available from: http://www.gov.nu.ca/sites/default/files/files/Lung_Bronchus_Final_16Jul2013(4).pdf.

103. Canadian Partnership Against Cancer. First Nations cancer control in Canada baseline report. Toronto, ON; 2013. Available from: http://www.cancerview.ca/idc/groups/public/documents/webcontent/first_nations_cc_baseline.pdf.

104.   Canadian Partnership Against Cancer. Examining disparities in cancer control. Toronto, ON; 2014. Available from: www.cancerview.ca/systemperformancereport

105.   Zhu C. Cancer screening in Aboriginal communities: a promising parctices review. Calgary, AB; 2012. Available from: http://www.albertahealthservices.ca/poph/hi-poph-aboriginal-health-review-2012.pdf.

106.   Kelly J, Lanier A, Santos M, Healey S, Louchini R, Friborg J, et al. Cancer among the circumpolar Inuit, 1989-2003. II. Patterns and trends. Int J Circumpolar Health. 2008; 67(5):408-20. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19186762.

107.   Canadian Partnership Against Cancer. First Nations, Inuit and Métis action plan on cancer control. Toronto, ON; 2011. Available from: http://www.partnershipagainstcancer.ca/wp-content/uploads/First-Nations-Inuit-and-M%C3%A9tis-Action-Plan-on-Cancer-Control-June-2011.pdf.

108.   Physicians for a Smoke-Free Canada [Internet]. Smoking among Aboriginal Canadians. Available from: http://www.smoke-free.ca/factsheets/pdf/cchs/aboriginal.pdf.

109.   Reid JL, Hammond D, Rynard VL, Burkhalter R. Tobacco use in Canada: patterns and trends. 2014 edition. Waterloo, ON: University of Waterloo; 2014. Available from: www.tobaccoreport.ca.

110.   Kubik AK, Parkin DM, Zatloukal P. Czech Study on Lung Cancer Screening: post-trial follow-up of lung cancer deaths up to year 15 since enrollment. Cancer. 2000; 89(11 Suppl):2363-8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11147613.

111.   Fukuoka M, Yano S, Giaccone G, Tamura T, Nakagawa K, Douillard JY, et al. Multi-institutional randomized phase II trial of gefitinib for previously treated patients with advanced non-small-cell lung cancer (The IDEAL 1 Trial) [corrected]. Journal of Clinical Oncology. 2003; 21(12):2237-46. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12748244.

112.   McMahon PM, Kong CY, Bouzan C, Weinstein MC, Cipriano LE, Tramontano AC, et al. Cost-effectiveness of computed tomography screening for lung cancer in the United States. Journal of Thoracic Oncology. 2011; 6(11):1841-8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21892105.

113.   Goulart BHL, Bensink ME, Mummy DG, Ramsey SD. Lung cancer screening with low-dose computed tomography: costs, national expenditures, and cost-effectiveness. J Natl Compr Cancer Netw. 2012; 10(2):267-75. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22308519.

114.   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. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23940744.

115.   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):347-55. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24418526.

116. Tota JE, Ramanakumar AV, Franco EL. Lung cancer screening: review and performance comparison under different risk scenarios. Lung. 2014; 192(1):55-63. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24153450.

117. Field JK, Chen Y, Marcus MW, McRonald FE, Raji OY, Duffy SW. The contribution of risk prediction models to early detection of lung cancer. J Surg Oncol. 2013; 108(5):304-11. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23996507.

118. Kubik A, Polak J. Lung cancer detection. Results of a randomized prospective study in Czechoslovakia. Cancer. 1986; 57(12):2427-37. Available from: http://www.ncbi.nlm.nih.gov/pubmed/3697941.

119. Walter SD, Kubik A, Parkin DM, Reissigova J, Adamec M, Khlat M. The natural history of lung cancer estimated from the results of a randomized trial of screening. Cancer Causes Control. 1992; 3(2):115-23. Available from: http://www.ncbi.nlm.nih.gov/pubmed/1562701.

120. Ashraf H, Tonnesen P, Holst Pedersen J, Dirksen A, Thorsen H, Dossing 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):388-92. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19052048.

121. Kaerlev L, Iachina M, Pedersen JH, Green A, Norgard BM. CT-Screening for lung cancer does not increase the use of anxiolytic or antidepressant medication. BMC Cancer. 2012; 12:188. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22621716.

122. 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. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22382119.

123. Pedersen JH, Petersen RH, Hansen HJ. Lung cancer screening trials: Denmark and beyond. The Journal of Thoracic and Cardiovascular Surgery. 2012; 144(3):S7-8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22513316.

124. Pedersen JH, Ashraf H, Dirksen A, Bach K, Hansen H, Toennesen P, 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):608-14. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19357536.

125. 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):347-55. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24418526.

126. Saghir Z, Ashraf H, Dirksen A, Brodersen J, Pedersen JH. Contamination during 4 years of annual CT screening in the Danish Lung Cancer Screening Trial (DLCST). Lung Cancer. 2011; 71(3):323-7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20619924.

127. Shaker SB, Dirksen A, Lo P, Skovgaard LT, De Bruijne M, Pedersen JH. Factors influencing the decline in lung density in a Danish lung cancer screening cohort. Eur Respir J. 2012; 40(5):1142-8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22408202.

128. Doria-Rose VP, Marcus PM, Szabo E, Tockman MS, Melamed MR, Prorok PC. Randomized controlled trials of the efficacy of lung cancer screening by sputum cytology revisited: a combined mortality analysis from the Johns Hopkins Lung Project and the Memorial Sloan-Kettering Lung Study. Cancer. 2009; 115(21):5007-17. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19637354.

129. Frost JK, Ball WC, Jr., Levin ML, Tockman MS, Baker RR, Carter D, et al. Early lung cancer detection: results of the initial (prevalence) radiologic and cytologic screening in the Johns Hopkins study. The American Review of Respiratory Disease. 1984; 130(4):549-54. Available from: http://www.ncbi.nlm.nih.gov/pubmed/6091505.

130. Levin ML, Tockman MS, Frost JK, Ball WC, Jr. Lung cancer mortality in males screened by chest X-ray and cytologic sputum examination: a preliminary report. Recent Results in Cancer Research. 1982; 82:138-46. Available from: http://www.ncbi.nlm.nih.gov/pubmed/7111836.

131. Tockman MS. Sensitivity of methacholine testing in occupational asthma. Chest. 1986; 89(Suppl 4):S324-5. Available from: http://journal.publications.chestnet.org/article.aspx?articleid=1058961.

132. Friedman GD, Collen MF, Fireman BH. Multiphasic Health Checkup Evaluation: a 16-year follow-up. J Chronic Dis. 1986; 39(6):453-63. Available from: http://www.ncbi.nlm.nih.gov/pubmed/3711252.

133. Flehinger BJ, Kimmel M, Polyak T, Melamed MR. Screening for lung-cancer - the Mayo Lung Project revisited. Cancer. 1993; 72(5):1573-80. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8394199.

134. Fontana RS, Sanderson DR, Woolner LB, Miller WE, Bernatz PE, Payne WS, et al. The Mayo Lung Project for early detection and localization of bronchogenic carcinoma: a status report. Chest. 1975; 67(5):511-22. Available from: http://www.ncbi.nlm.nih.gov/pubmed/1126186.

135. Marcus PM, Bergstralh EJ, Fagerstrom RM, Williams DE, Fontana R, Taylor WF, et al. Lung cancer mortality in the Mayo Lung Project: impact of extended follow-up. J Natl Cancer Inst. 2000; 92(16):1308-16. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10944552.

136. Marcus PM, Prorok PC. Reanalysis of the Mayo Lung Project data: the impact of confounding and effect modification. J Med Screen. 1999; 6(1):47-9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10321372.

137. Sanderson D, Fontana R. Results of Mayo lung project: an interim report. Recent Results Cancer Research. 1982; 82:179-86. Available from: http://www.ncbi.nlm.nih.gov/pubmed/6287546.

138. Sanderson DR. Lung cancer screening: The Mayo Study. Chest. 1986; 89(Suppl 4):S324. Available from: http://journal.publications.chestnet.org/article.aspx?articleid=1058959.

139. Shi L, Tian H, McCarthy WJ, Berman B, Wu S, Boer R. Exploring the uncertainties of early detection results: model-based interpretation of Mayo Lung Project. BMC Cancer. 2011; 11:92. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21375784.

140. Strauss GM. The Mayo Lung Cohort: a regression analysis focusing on lung cancer incidence and mortality. Journal of Clinical Oncology. 2002; 20(8):1973-83. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11956255.

141. Taylor WF, Fontana RS. Biometric design of the Mayo Lung Project for early detection and localization of bronchogenic carcinoma. Cancer. 1972; 30(5):1344-7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/5083070.

142. Taylor WF, Fontana RS, Uhlenhopp MA, Davis CS. Some results of screening for early lung cancer. Cancer. 1981; 47(5 Suppl):1114-20. Available from: http://www.ncbi.nlm.nih.gov/pubmed/6263442.

143. Woolner LB, Fontana RS, Cortese DA, Sanderson DR, Bernatz PE, Payne WS, et al. Roentgenographically occult lung cancer: pathologic findings and frequency of multicentricity during a 10-year period. Mayo Clin Proc. 1984; 59(7):453-66. Available from: http://www.ncbi.nlm.nih.gov/pubmed/6738113.

144. Flehinger BJ, Melamed MR. Current status of screening for lung cancer. Chest Surg Clin N Am. 1994; 4(1):1-15. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8055275.

145. Martini N. Results of the Memorial-Sloan-Kettering Study in screening for early lung-cancer. Chest. 1986; 89(Suppl 4):S325. Available from: http://journal.publications.chestnet.org/article.aspx?articleID=1058963.

146. Melamed M, Flehinger B, Miller D, Osborne R, Zaman M, Mcginnis C, et al. Preliminary-report of lung-cancer detection program in New-York. Cancer. 1977; 39(2):369-82. Available from: http://www.ncbi.nlm.nih.gov/pubmed/837325.

147. Aberle DR, Abtin F, Brown K. Computed tomography screening for lung cancer: has it finally arrived? Implications of the National Lung Screening Trial. Journal of Clinical Oncology. 2013; 31(8):1002-8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23401434.

148. Chiles C, Paul NS. Beyond lung cancer: A strategic approach to interpreting screening computed tomography scans on the basis of mortality data from the National Lung Screening Trial. J Thorac Imaging. 2013; 28(6):347-54. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24071622.

149. Hensing TA, Salgia R. Molecular biomarkers for future screening of lung cancer. J Surg Oncol. 2013; 108(5):327-33. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23893423.

150. Kovalchik SA, Tammemagi M, Berg CD, Caporaso NE, Riley TL, Korch M, et al. Targeting of low-dose CT screening according to the risk of lung-cancer death. N Engl J Medicine. 2013; 369(3):245-54. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23863051.

151. McCunney RJ, Li J. Radiation risks in lung cancer screening programs: a comparison with nuclear industry workers and atomic bomb survivors. Chest. 2014; 145(3):618-24. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24590022.

152. National Lung Screening Trial Research Team, Church TR, Black WC, Aberle DR, Berg CD, Clingan KL, et al. Results of initial low-dose computed tomographic screening for lung cancer. N Engl J Medicine. 2013; 368(21):1980-91. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23697514.

153. National Lung Screening Trial Research Team, Aberle DR, Berg CD, Black WC, Church TR, Fagerstrom RM, et al. The National Lung Screening Trial: overview and study design. Radiology. 2011; 258(1):243-53. Available from: http://www.ncbi.nlm.nih.gov/pubmed/?term=21045183.

154. Pinsky PF, Church TR, Izmirlian G, Kramer BS. The National Lung Screening Trial: results stratified by demographics, smoking history, and lung cancer histology. Cancer. 2013; 119(22):3976-83. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24037918.

155. Oudkerk M, Heuvelmans MA. Screening for lung cancer by imaging: the NELSON study. JBR-BTR. 2013; 96(3):163-6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23971173.

156. van den Bergh KAM, Essink-Bot M-L, Bunge EM, Scholten ET, Prokop M, van Iersel CA, et al. Impact of computed tomography screening for lung cancer on participants in a randomized controlled trial (NELSON trial). Cancer. 2008; 113(2):396-404. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18484588.

157. van den Bergh KA, Essink-Bot ML, Borsboom GJ, 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. The European Respiratory Journal. 2011; 38(1):154-61. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21148229.

158. van der Aalst CM, van Iersel CA, van Klaveren RJ, Frenken FJM, Fracheboud J, Otto SJ, et al. Generalisability of the results of the Dutch-Belgian randomised controlled lung cancer CT screening trial (NELSON): does self-selection play a role? Lung Cancer. 2012; 77(1):51-7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22459203.

159. van der Aalst CM, van den Bergh KA, Willemsen MC, de Koning HJ, van Klaveren RJ. Lung cancer screening and smoking abstinence: 2 year follow-up data from the Dutch-Belgian randomised controlled lung cancer screening trial. Thorax. 2010; 65(7):600-5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20627916.

160. van der Aalst CM, van Klaveren RJ, van den Bergh KA, Willemsen MC, de Koning HJ. The impact of a lung cancer computed tomography screening result on smoking abstinence. The European Respiratory Journal. 2011; 37(6):1466-73. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21148233.

161. Zhao YR, de Bock GH, Vliegenthart R, van Klaveren RJ, Wang Y, Bogoni L, et al. Performance of computer-aided detection of pulmonary nodules in low-dose CT: comparison with double reading by nodule volume. Eur Radiol. 2012; 22(10):2076-84. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22814824.

162. Zhao YR, Xie X, de Koning HJ, Mali WP, Vliegenthart R, Oudkerk M. NELSON lung cancer screening study. Cancer Imaging. 2011; 11 Spec No A:S79-84. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3266562/.

163. Brett GZ. The value of lung cancer detection by six-monthly chest radiographs. Thorax. 1968; 23(4):414-20. Available from: http://www.ncbi.nlm.nih.gov/pubmed/5664703.

164. Brett GZ. Earlier diagnosis and survival in lung cancer. Br Med J. 1969; 4(5678):260-2. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1629689/.

165. Brett GZ. The presymptomatic diagnosis of lung cancer. Proc R Soc Med. 1966; 59(11 Part 2):1208-14. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1900671/pdf/procrsmed00176-0043.pdf.

166. Kramer BS, Gohagan J, Prorok.P., National Cancer Institute DoCPaC. A randomized study of chest x-ray screening for lung cancer as part of the Prostate, Lung, Colorectal, and Ovarian (PLCO) Trial. Lung Cancer. 1994; 11(Suppl 2):S82-3. Available from: http://www.sciencedirect.com/science/journal/01695002/11/supp/S2.

167. Hocking WG, Hu P, Oken MM, Winslow SD, Kvale PA, Prorok PC, et al. Lung cancer screening in the randomized Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial. J Natl Cancer Inst. 2010; 102(10):722-31. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20442215.

168. Oken MM, Marcus PM, Hu P, Beck TM, Hocking W, Kvale PA, et al. Baseline chest radiograph for lung cancer detection in the randomized Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial. J Natl Cancer Inst. 2005; 97(24):1832-9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16368945.

169. Prorok PC, Andriole GL, Bresalier RS, Buys SS, Chia D, Crawford ED, et al. Design of the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. Control Clin Trials. 2000; 21(Suppl 6):273S-309S. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11189684.

170. Andriole Gl, Crawford ED, Grubb RL, Buys SS, Chia D, Church TR, et al. Prostate cancer screening in the randomized Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial: mortality results after 13 years of follow-up. J Natl Cancer Inst. 2012; 104(2):125-32. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22228146.

171. Croswell JM, Kramer BS, Kreimer AR, Prorok PC, Xu JL, Baker SG, et al. Cumulative incidence of false-positive results in repeated, multimodal cancer screening. Ann Fam Med. 2009; 7(3):212-22. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19433838.

172. MacRedmond R, Logan PM, Lee M, Kenny D, Foley C, Costello RW. Screening for lung cancer using low dose CT scanning. Thorax. 2004; 59(3):237-41. Available from: http://thorax.bmj.com/content/59/3/237.abstract.

173. Buys SS, Partridge E, Black A, Johnson CC, Lamerato L, Isaacs C, et al. Effect of screening on ovarian cancer mortality: the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer

Screening Randomized Controlled Trial. JAMA. 2011; 305(22):2295-303. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21642681.

174.   Schoen RE, Pinsky PF, Weissfeld JL, Yokochi LA, Church T, Laiyemo AO, et al. Colorectal-cancer incidence and mortality with screening flexible sigmoidoscopy. N Engl J Medicine. 2012; 366(25):2345-57. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22612596.

175.   Veronesi G, Bellomi M, Scanagatta P, Preda L, Rampinelli C, Guarize J, et al. Difficulties encountered managing nodules detected during a computed tomography lung cancer screening program. The Journal of Thoracic and Cardiovascular Surgery. 2008; 136(3):611-7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18805261.

176.   Sterne JAC, Egger M, Moher D, The Cochrane Bias Methods Group. Addressing reporting biases. In: J. P. T. Higgins, S. Green, editors. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]*: The Cochrane Collaboration; 2011.

Figures

Figure 1: Analytic Framework

Figure 2: Search and Selection Results

Figure 1: Analytic Framework

INTERVENTION
Lung cancer screening
(CXR, SC, LDCT)

POPULATION
Adults ≥18 years at
average and high
risk who are not
suspected of
having lung cancer

Key Question 1

Key Question 2

CLINICAL BENEFITS
Lung cancer mortality,
All-cause mortality,
Smoking cessation rate,
Stage at diagnosis,
Incidental findings

HARMS
Overdiagnosis,
Death from invasive follow-up testing
Major complications or morbidity from invasive follow-up testing,
False positives,
Consequences of false positives,
Negative consequences of incidental findings,
Anxiety,
Quality of life,
Infection from invasive follow-up testing,
Bleeding from invasive follow-up testing

Figure 2: Search and Selection Results

| USPSTF Review 16 | Cochrane Review 5 | Hand Searched Reference Lists 18 | Update Search for Benefits & Extended Search for Harms 840 |

Title & Abstract Screening 879

Excluded at Title & Abstract Screening 656

Eligible for Full Text Screening 223

Excluded at Full Text 161

Reasons for Exclusion:
Population: 63
Intervention: 17
Comparison: 45
Design: 34
Outcomes: 2

Systematic Reviews 29

Included Studies 33*

Benefits 13

Harms 30

*There are 33 included studies in total; 10 studies appear in both Key Questions

Tables

Table 1: Summary of Risk of Bias Assessment of RCTs Included for KQ1 (Benefits of Screening)

Table 2: Broad Features of the Available Evidence for KQ1 (Benefits of Screening)

Table 3: Characteristics of RCTs Included for KQ1 (Benefits of Screening)

Table 4: List of Studies Included for KQ2 (Harms of Screening or Follow-up Testing)

Table 5: Overall Findings Summary - Benefits (Critical and Selected Important Outcomes)

Table 6: Overall Findings Summary - Harms (Critical Outcomes)

Table 1: Summary of Risk of Bias Assessment of RCTs Included for KQ1 (Benefits of Screening)

Study	Sequence Generation	Allocation Concealment	Blinding of Outcome Assessment		Incomplete Reporting of Outcomes		Selective Reporting	Other**
			Objective*	Self-report*	Objective*	Self-report*		
Czech Study[42]	L	U	U	U	U	U	L	L
DANTE[21]	L	U	H		H		L	H
DLCST[40]	L	U	U	U	L	L	L	U
Johns Hopkins[43]	L	U	L		L		L	L
Kaiser Foundation[44]	H	U	L		H		L	H
Lung Screening Study[41]	L	U	U		L		L	L
Mayo Lung Project[39]	U	U	L	U	L	L	L	L
Memorial Sloan-Kettering[45]	L	L	L		L		L	L
MILD[23]	U	U	L		L		L	H
NELSON[37]	L	U		U		L	L	L
NLST[10]	L	U	L		L		L	L
North London Study[46]	L	U	U		L		L	L
PLCO[38]	L	U	L		L		L	L

L (green) = Low Risk; U (yellow) = Unclear Risk; H (red) = High Risk

*objectively assessed outcomes: lung cancer mortality, all-cause mortality, stage at diagnosis, incidental findings; self-report outcome: smoking cessation

**other potential sources of bias: industry funding; study underpowered or small sample (<30 participants per arm); control group contamination through opportunistic screening; baseline differences

Table 2: Broad Features of the Available Evidence for KQ1 (Benefits of Screening)

Designs	• 13 RCTs
Populations	• 7 studies had a mixed gender sample; 6 studies included men only • 10 studies included current or former smokers; 2 studies included smokers and non-smokers; 1 study included never smokers, former smokers and current smokers • studies targeted adults 35 to 74 years; 6 studies recruited some participants aged <50 years, 9 studies included some participants aged >70 years
Interventions	• 2 studies compared CXR to usual care • 5 studies compared intensive CXR (with or without SC) to less-intensive screening • 4 studies compared LDCT to no screening/usual care • 2 studies compared LDCT to CXR
Length of Follow-up	• 10 studies had follow-up ≥5 years • 3 studies had follow-up <5 years
Quality Assessment	• 7 studies had unclear risk of bias • 3 studies were rated as having low risk of bias • 3 studies had high risk of bias
Study Locations	• 7 studies were conducted in the US • 6 studies were conducted in Europe (Denmark, Belgium, the Netherlands, Italy, Czechoslovakia)
Start Dates	• 6 studies initiated recruitment after 2000 • 1 study began in the 1990s • 4 studies were initiated in the 1970s • 2 studies were started in the 1960s

Table 3: Characteristics of RCTs Included for KQ1 (Benefits of Screening)

Study, Start Date, Location	Czech Study,[42] 1976, Czechoslovakia Companion Papers: Kubik,[71, 110, 118] Walter[119]
Objective	To evaluate semi-annual screening by radiology and sputum cytology in comparison to screening at a 3-year interval and to no screening
Methods	Design: RCT
	Recruitment: selected from general population of 6 districts in Czechoslovakia; subjects who attended the chest clinic from June 1, 1976 to June 30, 1977 and met the high risk criteria were referred to initial screening
	Inclusion Criteria: males aged 40 to 64 years with lifetime consumption of ≥150,000 cigarettes; current smoker; no known pulmonary disease
	Exclusion Criteria: unlikely to participate in periodic screenings for ≥3 years
Participants	Sample: 6,345
	Intervention n=3,171; Control n=3,174
	Age: range for inclusion 40 to 64 years; at initial screening: 40 to 44 years 15%, 45 to 49 years 22%, 50 to 54 years 29%, 55 to 59 years 18%, 60 to 64 years 14%
	Gender: 100% male
	Race/Ethnicity: not reported
	Smoking Status/History: all current and heavy smokers; 31 to 32 mean years of smoking
	Loss to Follow-up: not reported
Intervention	Description of Intervention: initial CXR and SC exams; after randomization 6 exams at 6-month intervals involving posteroanterior CXR with double independent reading and a simple type of SC investigation; over the course of an additional 3 years of follow-up, a CXR was done each year
	Description of Control: initial CXR and SC; 3 years after randomization control participants underwent 1 exam by the same 2 methods; over the course of an additional 3 years of follow-up, a CXR was done each year
	Screening Phase: 6 years (6 exams at 6 month intervals over 3 years) plus three annual CXR screens in years 4 to 6
	Length of Follow-up: 6 years
Study, Start Date, Location	DANTE,[21] 2001, Italy
Objective	To explore the effect of screening with LDCT on lung cancer mortality as well as on incidence, stage at diagnosis and resection
Methods	Design: RCT
	Recruitment: research was conducted by the Istituto Clinico Humanitas

	Hospital, Milan, Italy; accrual March 2001 to February 2006; two centers of the same hospital network, the Humanitas Gavazzeni Hospital in Bergamo and the Humanitas Oncology Center in Catania, only enrolled subjects for the trial during the last year of accrual; participants were assessed and randomized at a telephone interview with the project assistant
	Inclusion Criteria: male smokers or former smokers with ≥20 pack-years and aged 60 to 74 years
	Exclusion Criteria: comorbid conditions carrying a life expectancy <5 years; history of previous malignancy treated in last 10 years (a 5-year disease free interval was acceptable for early laryngeal cancer and non-melanoma skin cancer); unable to comply with protocol
Participants	Sample: 2,472
	Intervention n=1,276; Control n=1,196
	Age [Mean (Range)]: Intervention: 64.3 (64.0 to 64.7) years; Control: 64.6 (64.3 to 64.9) years
	Gender: 100% male
	Race/Ethnicity: not reported
	Smoking Status/History: current and former smokers; mean pack-years Intervention: 47.3; Control: 47.2
	Loss to Follow-up: not reported
Intervention	Description of Intervention: 5 annual LDCT screening rounds (baseline and 4 repeats) along with a medical interview and physical exam that focused on reassessing smoking habits, recent medical history, and new signs and symptoms of possible neoplasia since last assessment
	Description of Control: same annual medical interview and physical exam as intervention; no further evaluation in absence of clinically detected abnormalities
	Screening Phase: 5 years (one screen at baseline and 4 annually)
	Length of Follow-up: median 33.7 months (range 1.8 to 79.2 months)
Study, Start Date, Location	DLCST,[40] 2004, Denmark Companion Papers: Ashraf,[120] Kaerlev,[121] Mosborg,[122] Pedersen,[123, 124] Petersen,[72] Rasmussen,[125] Saghir,[19, 126] Shaker[127]
Objective	To investigate smoking behaviour of participants undergoing a complete 5-year screening program
Methods	Design: RCT
	Recruitment: ads for volunteers in free local and regional newspapers
	Inclusion Criteria: current or former smokers aged 50 to 70 years with a smoking history of >20 pack-years; former smokers had to have quit after age 50, and <10 years prior
Participants	Sample: 4,104

	Intervention n=2,052; Control n=2,052
	Age [Mean (Range)]: Intervention: 57.9 (49 to 71) years; Control: 57.8 (49 to 71) years
	Gender (Male): Overall: 55%; Intervention n=1,147; Control n=1,120
	Race/Ethnicity: not reported
	Smoking Status/History: current or former (quit after age 50 and <10 years prior to enrollment) smokers with ≥20 pack-years (mean 36 pack-years), mean consumption of 19 cigarettes per day
	Loss to Follow-up: Intervention n=15; Control n=14
Intervention	Description of Intervention: 5 annual CT scans of the thorax; annual visits where smoking status was determined along with other assessments
	Description of Control: no screening; annual visits where smoking status was determined along with other assessments
	Screening Phase: 5 years (4 years of annual screening)
	Length of Follow-up: 5 years
Study, Start Date, Location	Johns Hopkins Study,[43] 1973, US Companion papers: Doria-Rose,[128] Frost,[129] Levin,[130] Tockman[131]
Objective	To determine whether the combination of CXR and SC would result in a reduction of lung cancer mortality in comparison to screening by CXR alone
Methods	Design: RCT
	Recruitment: via mail-outs and local industrial and occupation groups in Baltimore; majority of participants recruited via direct mailings to licensed male drivers ≥45 years (mailing list compiled by the Department of Motor Vehicles of the State of Maryland)
	Inclusion Criteria: currently (or within last year) smoke ≥1 pack/day
Participants	Sample: 10,387
	Intervention n=5,161; Control n=5,226
	Age: for inclusion ≥45 years; enrolled: 45 to 49 years 31%, 50 to 54 years 27%, 55 to 59 years 25%, 60 to 64 years 13%, 65 to 69 years 7%, 70+ years 2 to 3%
	Gender: 100% male
	Race/Ethnicity: not reported
	Smoking Status/History: all smokers (≥1 pack/day); pack-years: <25 6%, 25 to 49 41%, 50 to 74 33%, 75+ 20%
	Loss to Follow-up: not reported
Intervention	Description of Intervention 1: annual CXR for 5 to 7 years and SC every 4 months for 5 to 7 years
	Description of Intervention 2: annual CXR for 5 to 7 years

	Description of Control: no control group
	Screening Phase: 5 to 7 years (annual CXR and SC every 4 months)
	Length of Follow-up: 9 years
Study, Start Date, Location	Kaiser Foundation Study,[44] 1964, US Companion Paper: Friedman[132]
Objective	To assess the impact of attempting to increase adults' participation in periodic health exams on morbidity, disability and mortality
Methods	Design: RCT
	Recruitment: in 1964, a group of 46,000 Kaiser Foundation Health Plan members was selected as a source of subjects for the study; a sampling method of picking certain digits in medical record numbers was used to select 2 groups, each with at least 5,000 members
	Inclusion Criteria: men and women aged 35 to 54 (born 1910 to 1929); lived in the San Francisco Bay Area; Health Plan members for ≥2 years
Participants	Sample: 10,713
	Intervention n=5,156; Control n=5,557
	Age: range for inclusion 35 to 54 years; enrolled: 52% aged 45 to 54, 48% aged 35 to 44
	Gender (Male): Intervention: n=2,365 (45.9%); Control: n=2,643 (47.6%)
	Race/Ethnicity (Black or Asian): Intervention: 30.7%; Control: 30%
	Smoking Status/History: smokers (17%) and non-smokers
	Loss to Follow-up: not reported
Intervention	Description of Intervention: Kaiser Permanente Multiphasic Health Checkup (MHC) consisted of: (a) lab visit to complete questionnaire and tests including electrocardiography, sphygmomanometry, audiometry, visual acuity and tonometry, spirometry, CXR, mammography for women ≥48 years, urinalysis, and blood tests; (b) gynecologic exam and Pap smear for all women and sigmoidoscopic exam for all persons >40 years; and (c) follow-up visit with physician (or nurse practitioner) for physical exam and review of results
	Description of Control: no encouragement to have MHC but provided if requested
	Duration of Intervention: 16 years of annually available screening
	Length of Follow-up: 16 years
Study, Start Date, Location	Lung Screening Study,[41] 2000, US
Objective	To determine whether scanning with LDCT can reduce lung cancer mortality
Methods	Design: RCT
	Recruitment: 6 Lung Screening Study centres mailed 653,417 information packages in September 2000; recruitment primarily through mass mailings, but

	also included posters, advertisements, and recommendations from practitioners
	Inclusion Criteria: aged 55 to 74 years; current or a former (quit in last 10 years) cigarette smoker; ≥30-pack-year smoking history
	Exclusion Criteria: spiral CT scan of lungs or thorax in past 24 months; history of lung cancer; receiving treatment for any cancer except non-melanoma skin cancer; a portion of a lung or an entire lung removed; enrolled in another cancer screening trial (including the PLCO study) or primary cancer prevention trial other than for smoking cessation
Participants	Sample: 3,318
	Intervention 1 n=1,660; Intervention 2 n=1,658
	Age: range for inclusion 55 to 74 years; enrolled: 55 to 59 years 37%, 60 to 64 years 30%, 65 to 69 years 21%, 70 to 74 years 11%
	Gender (Male): Intervention 1: n=965 (58.1%); Intervention 2: n=978 (59.0%)
	Race/Ethnicity: not reported
	Smoking Status/History: former (quit <10 years; 57%) and current smokers (43%)
	Loss to Follow-up: Intervention n=15; Control n=14
Intervention	Description of Intervention 1: 1 LDCT
	Description of Intervention 2: 1 CXR
	Description of Control: no control group
	Screening Phase: 1 screen
	Length of Follow-up: 1 year
Study, Start Date, Location	Mayo Lung Project,[39] 1971, US Companion Papers: Flehinger,[133] Fontana,[70, 132, 134] Marcus,[135, 136] Sanderson,[137, 138] Shi,[139] Strauss,[140] Taylor,[141, 142] Woolner,[143] Yankelevitz[66]
Objective	To assess the effect of an intense regimen of CXR and SC on lung cancer mortality
Methods	Design: RCT
	Recruitment: recruitment of Mayo Clinic outpatients; enrolled 9,211 older male smokers who tested negative for lung cancer on CXR and SC screening
	Inclusion Criteria: men aged ≥45 years; life expectancy ≥5 years; sufficient respiratory reserve to undergo lobectomy if necessary
Participants	Sample: 9,211
	Intervention n=4,618; Control n=4,593
	Age: range for inclusion ≥45 years; enrolled: <50 years 25%, 50 to 54 years 24%, 55 to 59 years 22%, 60 to 64 years 17%, 65 to 69 years 10%, 70+ years 1%
	Gender: 100% male
	Race/Ethnicity: not reported

	Smoking Status/History: 94% smoked ≥1 packs a day; 97% had smoked ≥20 years; 78% had smoked ≥30 years; 34% had smoked ≥40 years; 91% had 25 to 49 pack-years; 43% had ≥50 pack-years
	Loss to Follow-up: Overall n=26
Intervention	Description of Intervention: CXR and SC tests every 4 months for 6 years
	Description of Control: advised to have yearly CXR and SC tests
	Screening Phase: 6 years (CXR and SC every 4 months)
	Length of Follow-up: 20.5 years
Study, Start Date, Location	Memorial Sloan-Kettering Study,[45] 1974, US Companion Papers: Flehinger,[144] Martini,[145] Melamed,[69, 146] Yankelevitz[66]
Objective	To investigate whether SC every 4 months and a yearly CXR leads to earlier detection of lung cancer; whether early, cytologically detected cancers can be consistently localized and treated; and whether this type of screening has any impact on lung cancer mortality
Methods	Design: RCT
	Recruitment: participants recruited through Group Health Incorporated, New York City police, motor vehicle registrants, newspaper notices, television, radio, companies/unions, and word of mouth
	Inclusion Criteria: men aged ≥45 years
	Exclusion Criteria: history of lung cancer
Participants	Sample: 10,040
	Intervention 1 n=5,072; Intervention 2 n=4,968
	Age: range for inclusion ≥45 years; enrolled: 45 to 49 years 34%, 50 to 54 years 26%, 55 to 59 years 18%, 60 to 64 years 13%, 65 to 69 years 7%, 70+ years 3%
	Gender: 100% male
	Race/Ethnicity: not reported
	Smoking Status/History: all smokers ≥1 pack/day; could have quit within the last year; pack-years <25 9%, 25 to 49 40%, 50 to 74 33%, 75+ 18%
	Loss to Follow-up: not reported
Intervention	Description of Intervention 1: annual CXR and SC every 4 months
	Description of Intervention 2: annual CXR
	Screening Phase: 5 years (CXR and SC every 4 months)
	Length of Follow-up: 9 years
Study, Start Date, Location	MILD,[23] 2005, Italy
Objective	To evaluate the effect of annual or biennial LDCT for early lung cancer detection on mortality

Methods	Design: RCT
	Recruitment: respondents to advertisements seeking volunteers published in newspapers and on television; enrolled from September 2005 to January 2011
	Inclusion Criteria: aged ≥49 years; current or former (quit in last 10 years) smokers with ≥20 pack-years; no history of cancer in past 5 years
Participants	Sample: 4,099
	Intervention 1 n=1,190; Intervention 2 n=1,186; Control n=1,723
	Age (Median, Years): Intervention 1=57; Intervention 2=58; Control=57
	Gender (Male): Intervention 1: n=814 (68.4%); Intervention 2: n=813 (68.5%); Control: n=1,090 (63.3%)
	Race/Ethnicity: not reported
	Smoking Status/History: current (Intervention 69%; Control 90%) and former (Intervention 32%; Control 10%) smokers with ≥20 years smoking duration; mean pack-years 38 to 39
	Loss to Follow-up: not reported
Intervention	Description of Intervention 1: annual screening with LDCT, as well as a smoking cessation program
	Description of Intervention 2: biennial screening with LDCT, as well as smoking cessation program
	Description of Control: smoking cessation program
	Screening Phase: 5 years of annual or biennial screening
	Length of Follow-up: 5 years
Study, Start Date, Location	NLST,[10] 2002, US Companion Papers: Aberle,[82, 147] Chiles,[148] Duffy,[65] Hensing,[149] Kovalchik,[150] McCunney,[151] NLSTRT,[152, 153] Patz,[67] Pinsky[154]
Objective	To determine the effect of screening with LDCT on mortality from lung cancer
Methods	Design: RCT
	Recruitment: high risk persons enrolled at 33 US medical centers
	Inclusion Criteria: men and women; aged 55 to 74 years; history of cigarette smoking ≥30 pack-years; former smokers who quit <15 years; asymptomatic
	Exclusion Criteria: recent diagnosis of lung cancer; chest CT within last 18 months; hemoptysis; unexplained weight loss >6.8 kg in last year
Participants	Sample: 53,454
	Intervention 1 n=26,722; Intervention 2 n=26,732
	Age: range for inclusion 55 to 74 years; enrolled 55 to 59 years 47%, 60 to 64 years 31%, 65 to 69 years 18%, 70 to 74 years 9%
	Gender (Male): Intervention 1: n=15,770 (59%); Intervention 2: n=15,762 (59%)

	Race/Ethnicity: Caucasian - Intervention 1: n=24,289 (90.9%), Intervention 2: n=24,260 (90.8%); Black - Intervention 1: n=1,195 (4.5%), Intervention 2: n=1,181 (4.4%); Other - Intervention 1: n=1,238 (4.6%), Intervention 2: n=1,291 (4.8%)
	Smoking Status/History: current (48%) and former (quit <15 years; 52%) smokers
	Loss to Follow-up: not reported
Intervention	Description of Intervention 1: 3 annual screens using LDCT
	Description of Intervention 2: 3 annual screens using CXR
	Screening Phase: 3 years (annual screening)
	Length of Follow-up: 6.5 years (median)
Study, Start Date, Location	NELSON,[37] 2003, Netherlands, Belgium Companion Papers: Horeweg,[75] Oudkerk,[155] van den Bergh,[79, 156, 157] van der Aalst,[158-160] van Iersel,[18] Zhao[161, 162]
Objective	To study the impact of strict referral criteria and increasing screening interval on characteristics of screen-detected lung cancers and to compare the findings across screening rounds, between genders and with other screening trials
Methods	Design: RCT
	Recruitment: in 2003 addresses of all men born in 1928 to 1952 were obtained from population registries in 7 districts in the Netherlands; addresses of all men and women of the same age were obtained from population registries of 14 municipalities in Belgium; of 335,441 individuals who were mailed a questionnaire about health, lifestyle and smoking history, respondents who met the eligibility criteria received an invitation for screening, an information leaflet, and an informed consent form with a short questionnaire
	Inclusion Criteria: men and women; aged 50 to 75 years; former and current smokers; adequate health status
	Exclusion Criteria: unable to climb 2 flights of stairs and body weight ≥140 kg; current or past renal cancer, melanoma or breast cancer; lung cancer diagnosed <5 years ago; lung cancer diagnosed ≥5 years ago but still under treatment; chest CT exam <1 year before completing first NELSON questionnaire
Participants	Sample: 15,822
	Intervention n=7,915; Control n=7,907
	Age [Mean (SD)]: Intervention: 57.8 (5.5) years; Control: 57.8 (5.7) years
	Gender (Male): n=6,328 (83.5%)
	Race/Ethnicity: not reported
	Smoking Status/History: current (56%) or former (quit <10 years; 44%) smokers; median 38 pack-years
	Loss to Follow-up: Intervention n=922; Control n=not reported

Intervention	Description of Intervention: screening with LDCT at baseline (first round), 1 year later (second round), 3 years post baseline (third round), and 5.5 years post baseline (fourth round)
	Description of Control: no screening
	Screening Phase: up to 5.5 years (1 to 4 screens)
	Length of Follow-up: 10 years
Study, Start Date, Location	North London Study,[46] 1960, UK Companion Papers: Brett[163-165]
Objective	To evaluate early lung cancer detection by CXR every 6 months
Methods	Design: RCT
	Recruitment: industrial workplaces in Northwest London were assessed for the feasibility of mobile x-ray visits every 6 months; firms were categorized according to type of work and location, and randomized into test and control groups; all sites were visited to explain study to management and workers
	Inclusion Criteria: men aged ≥40 years; smokers and non-smokers; employees of participating industrial workplaces in Northwest London
Participants	Sample: 55,034
	Intervention n=29,723; Control n=25,311
	Age: range for inclusion ≥40 years; enrolled: 50 to 59 years 47%, 60 to 64 years 31%, 65 to 69 years 18%, 70 to 74 years 9%
	Gender: 100% male
	Race/Ethnicity: not reported
	Smoking Status/History: non-smokers (12%), former smokers (19%) and smokers (69%)
	Loss to Follow-up: not reported
Intervention	Description of Intervention 1: 1 baseline CXR, followed by 1 CXR every 6 months for 3 years for a total of 7 screens
	Description of Intervention 2: 2 CXRs: 1 at baseline and 1 at end of study period
	Screening Phase: 3 years (7 screens)
	Length of Follow-up: 2 years
Study, Start Date, Location	PLCO,[38] 1993, US Companion Papers: Barnett,[166] Hocking,[77, 167] Lacasse,[81] Oken,[168] Prorok,[169] Taylor,[80] Andriole[170]
Objective	To evaluate the effect on mortality of screening for lung cancer using radiographs in the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial
Methods	Design: RCT
	Recruitment: enrollment initiated in 1993 (completed in 2001) at 10 screening

	centers across the US; mass mailing used to target residents from the general population in the catchment of each center
	Inclusion Criteria: men and women aged 55 to 74 years; smokers, former smokers and non/never-smokers
	Exclusion Criteria: history of a PLCO cancer; current cancer treatment; 1 lung removed
Participants	Sample: 154,901
	Intervention n=77,445; Control n=77,456
	Age: range for inclusion 55 to 74 years; enrolled: 55 to 59 years 33.4%, 60 to 64 years 30.7%, 65 to 69 years 22.5/22.6%, 70 to 74 years 13.4%
	Gender (Male): Intervention: n=38,340 (49.5%); Control: n=38,345 (49.5%)
	Race/Ethnicity: White non-Hispanic - Intervention: n=66,874 (86.4%); Control: n=65,708 (84.8%); Black non-Hispanic - Intervention: n=3,883 (5.0%); Control: n=3,825 (4.9%); Hispanic - Intervention: n=1,421 (1.8%); Control: n=1,397 (1.8%); Asian - Intervention: n=2,791 (3.6%); Control: n=2,785 (3.6%); Other/Unknown - Intervention: n=2,476 (3.2%); Control: n=3,741 (4.8%)
	Smoking Status/History: never smokers (45%); former smokers (42%) and smokers (10%)
	Loss to Follow-up: Intervention n=36,042; Control: not reported
Intervention	Description of Intervention: CXR at baseline and then 3 additional annual CXR tests (4 screens over 4 years)
	Description of Control: usual care (no organized screening)
	Screening Phase: 3 years (annual)
	Length of Follow-up: 13 years

Table 4: List of Studies Included for KQ2 (Harms of Screening or Invasive Follow-up Testing)

Trial or First Author Name	Companion Papers
Becker[49]	
Blanchon[59]	
Byrne[55]	Wilson[78]
Callol[60]	
Crestanello[53]	Swensen[76]
Croswell[51]	Croswell[171]
Czech study (Kubik)[42]	Kubik,[71, 110, 118] Walter[119]
Diederich[63]	
DLCST study (Ashraf)[40]	Ashraf,[120] Kaerlev,[121] Mosborg,[122] Pedersen,[123, 124] Petersen,[72] Rasmussen,[125] Saghir,[19, 126] Shaker[127]
Dominioni[50]	
Duffy[65]	
Erfurt study (Wilde)[64]	
Henschke[62]	
Infante[21]	
ITALUNG study (Lopes Pegna)[48]	Lopes Pegna[20]
Lung Screening Study (Gohagan)[41]	
MacRedmond[58]	MacRedmond[172]
Mayo Lung Project (Marcus)[39]	Flehinger,[133] Fontana,[70, 132, 134] Marcus,[135, 136] Sanderson,[137, 138] Strauss,[140] Shi,[139] Taylor,[141, 142] Woolner,[143] Yankelevitz[66]
Mazzone[47]	
Memorial Sloan-Kettering study (Flehinger)[45]	Flehinger,[144] Martini,[145] Melamed,[69, 146] Yankelevitz[66]
Menezes[56]	
MILD study (Pastorino)[23]	
NELSON study (Horeweg)[37]	Horeweg,[75] Oudkerk,[155] van den Bergh,[79, 156, 157] van der Aalst,[158-160] van Iersel,[18] Zhao[161, 162]
NLST study[10]	Aberle,[82, 147] Chiles,[148] de Koning,[93] Duffy,[65] Hensing,[149] Kovalchik,[150] McCunney,[151] NLST Research Team,[152, 153] Patz,[67] Pinsky[154]
Pastorino[24]	
PLCO study (Oken)[38]	Andriole,[170] Barnett,[166] Buys,[173] Hocking,[77, 167] Lacasse,[81] Oken,[168] Prorok,[169] Schoen,[174] Taylor[80]
Rzyman[54]	
Sobue[61]	
Sone[52]	
Veronesi[57]	Veronesi[68, 175]

Table 5: Overall Findings Summary - Benefits (Critical and Selected Important Outcomes)

Outcome	CXR vs Usual Care	CXR vs CXR	CXR plus SC vs CXR	Annual LDCT vs Usual Care	Biennial LDCT vs Usual Care	LDCT vs CXR
Lung Cancer Mortality	RR 0.99 95% CI 0.92, 1.07 I^2=0%	RR 1.03 95% CI 0.74, 1.42 I^2=na	RR 1.01 95% CI 0.87, 1.18 I^2=59%	RR 1.35 95% CI 0.79, 2.29 I^2=32%	RR 1.25 95% CI 0.42, 3.70 I^2=na	RR 0.80 95% CI 0.70, 0.92 I^2=na absolute value per million 3,250 fewer, range from 1,271 fewer to 4,972 fewer ARR 0.33% NNS 308 (95% CI 201, 787)
All-Cause Mortality	RR 0.98 95% CI 0.96, 1.00 I^2=0%	-	RR 1.04 95% CI 0.97, 1.11 I^2=37%	RR 1.42 95% CI 0.91, 2.22 I^2=67%	RR 1.45 95% CI 0.79, 2.69 I^2=na	RR 0.94 95% CI 0.88, 1.00 I^2=na absolute value per million 4,571 fewer, range from 180 fewer to 8,709 fewer ARR 0.46% NNS 219 (95% CI 115, 5,556)
Stage at Diagnosis (Early Stage)	RR 1.14 95% CI 1.03, 1.25 I^2=na	-	RR 1.15 95% CI 0.98, 1.36 I^2=50%	RR 1.59 95% CI 1.11, 2.28 I^2=0%	-	RR 1.46 95% CI 1.33, 1.61 I^2=na
Stage at Diagnosis (Late Stage)	RR 0.93 95% CI 0.87, 0.98 I^2=na	-	RR 0.85 95% CI 0.75, 0.96 I^2=13%	RR 0.59 95% CI 0.43, 0.83 I^2=0%	-	RR 0.71 95% CI 0.65, 0.77 I^2=na

ARR=Absolute Risk Reduction; NNS=Number Needed to Screen

Table 6: Overall Findings Summary - Harms (Critical Outcomes)

Outcomes	CXR	CXR plus SC	LDCT
Overdiagnosis	–	TVDT >400 days: 2.27% to 6.98% of all cases of lung cancer diagnosed in the screened population were overdiagnosed TVDT >300 days: 4.55% to 16.28% of all cases of lung cancer diagnosed in the screened population were overdiagnosed	10.99% to 25.83% of all cases of lung cancer diagnosed in the screened population were overdiagnosed
Death from Invasive Follow-up Testing	28.60 deaths (95% CI 16.02, 41.17) per 1,000 patients undergoing invasive follow-up testing	47.67 deaths (95% CI 23.86, 71.49) per 1,000 patients undergoing invasive follow-up testing	11.18 deaths (95% CI 5.07, 17.28) per 1,000 patients undergoing invasive follow-up testing
Major Complications from Invasive Follow-up Testing	63.32 major complications (95% CI 42.92, 92.49) per 1,000 patients undergoing invasive follow-up testing	–	43.29 major complications (95% CI 32.00, 54.58) per 1,000 patients undergoing invasive follow-up testing

TVDT = tumor volume doubling time

Evidence Set 1: KQ1 Benefits of Screening – Lung Cancer Mortality

- ES Table 1.1: GRADE Evidence Profile - Effect of Lung Cancer Screening on Lung Cancer Mortality

- ES Table 1.2: GRADE Summary of Findings - Effect of Lung Cancer Screening on Lung Cancer Mortality

- Forest Plot 1.1: Effect of Lung Cancer Screening Using CXR on Lung Cancer Mortality

- Forest Plot 1.2: Effect of Lung Cancer Screening Using LDCT on Lung Cancer Mortality

ES Table 1.1: GRADE Evidence Profile - Effect of Lung Cancer Screening on Lung Cancer Mortality *

No. of Studies	Quality Assessment						No. of Participants		Effect				GRADE Quality Rating	Importance
	Design	Risk of Bias	Inconsistency	Indirectness	Imprecision	Other	Screening	Control	Relative (95% CI)	Absolute per Million (Range)	ARR	NNS (95% CI)		
Lung Cancer Mortality: CXR screening vs. usual care (follow-up 10 to 16 years; assessed with: national or state death registries, death certificates[1])														
2[2]	randomized trials	serious risk[3]	no serious inconsistency[4]	no serious indirectness[5]	serious imprecision[6]	none[7]	1,257/82,583 (1.5221%)	1,272/82,992 (1.5327%)	RR 0.9908 (0.9171 to 1.0705)	141 fewer (1,271 fewer to 1,081 more)	-	-	⊕⊕OO LOW	CRITICAL
Lung Cancer Mortality: more intensive CXR screening vs. less intensive screening (follow-up 5 years; assessed with: hospital records, national death registry[1])														
1[8]	randomized trial	serious risk[9]	no serious inconsistency[10]	no serious indirectness[11]	serious imprecision[12]	none[7]	82/29,723 (0.2759%)	68/25,311 (0.2687%)	RR 1.0269 (0.7449 to 1.4156)	72 more (685 fewer to 1,117 more)	-	-	⊕⊕OO LOW	CRITICAL
Lung Cancer Mortality: more intensive CXR + SC screening vs. less intensive screening (follow-up 9 to 20.5 years; assessed with: national death registry, autopsies, death certificates, death-related clinical re cords[1])														
4[13]	randomized trials	serious risk[14]	no serious inconsistency[15]	no serious indirectness[16]	serious imprecision[17]	none[7]	840/17,983 (4.6711%)	812/18,000 (4.5111%)	RR 1.0141 (0.8717 to 1.1798)	636 more (5,788 fewer to 8,111 more)	-	-	⊕⊕OO LOW	CRITICAL
Lung Cancer Mortality: annual LDCT screening vs. usual care (follow-up 2.8 to 6 years; assessed with: national death registries, medical records[1])														
3[18]	randomized trials	serious risk[19]	no serious inconsistency[20]	no serious indirectness[21]	serious imprecision[22]	none[7]	47/4,518 (1.0403%)	38/4,971 (0.7644%)	RR 1.3460 (0.7904 to 2.2922)	2,645 more (1,602 fewer to 9,878 more)	-	-	⊕⊕OO LOW	CRITICAL
Lung Cancer Mortality: biennial LDCT screening vs. usual care (follow-up median 4.4 years; assessed with: national death registry[1])														
1[23]	randomized trial	serious risk[24]	no serious inconsistency[10]	no serious indirectness[25]	serious imprecision[26]	none[7]	6/1,186 (0.5059%)	7/1,723 (0.4063%)	RR 1.2452 (0.4195 to 3.6960)	996 more (2,358 fewer to 10,953 more)	-	-	⊕⊕OO LOW	CRITICAL
Lung Cancer Mortality: LDCT screening vs. CXR screening (follow-up median 6.5 years; assessed with: national death registry, death certificates, adjudicated cause[1])														
1[27]	randomized trial	no serious risk[28]	no serious inconsistency[10]	no serious indirectness[29]	no serious imprecision[30]	none[7]	356/26,722 (1.3322%)	443/26,732 (1.6572%)	RR 0.8039 (0.7000 to 0.9233)	3,250 fewer (1,271 fewer to 4,972 fewer)	0.33%	308 (201, 787)	⊕⊕⊕⊕ HIGH	CRITICAL

*Footnotes appear below the Summary of Findings Table

ES Table 1.2: GRADE Summary of Findings - Effect of Lung Cancer Screening on Lung Cancer Mortality

Outcome: Lung Cancer Mortality	Illustrative Comparative Risks* (95% CI)		Relative Effect (95% CI)	No. of Participants (Studies)	Quality of the Evidence (GRADE)
	Assumed Risk Number per Million Control	Corresponding Risk Number per Million Treatment			
CXR screening vs usual care (national or state death registries, death certificates[1]; follow-up: 10 to 16 years)	**15,327**	**15,186** (14,056 to 16,407)	**RR 0.9908** (0.9171 to 1.0705)	165,575 (2 studies[2])	⊕⊕⊖⊖ **low**[3,4,5,6,7]
more intensive CXR screening vs less intensive screening (hospital records, national death registry[1]; follow-up: 5 years)	**2,687**	**2,759** (2,001 to 3,803)	**RR 1.0269** (0.7449 to 1.4156)	55,034 (1 study[8])	⊕⊕⊖⊖ **low**[7,9,10,11,12]
more intensive CXR + SC screening vs. less intensive screening (national death registry, autopsies, death certificates, death-related clinical records[1]; follow-up: 9 to 20.5 years)	**45,111**	**45,747** (39,323 to 53,222)	**RR 1.0141** (0.8717 to 1.1798)	35,983 (4 studies[13])	⊕⊕⊖⊖ **low**[7,14,15,16,17]
annual LDCT screening vs usual care (national death registries, medical records[1]; follow-up: 2.8 to 6 years)	**7,644**	**10,289** (6,042 to 17,522)	**RR 1.3460** (0.7904 to 2.2922)	9,489 (3 studies[18])	⊕⊕⊖⊖ **low**[7,19,20,21,22]
biennial LDCT screening vs. usual care (national death registry[1]; follow-up: median 4.4 years)	**4,063**	**5,059** (1,704 to 15,016)	**RR 1.2452** (0.4195 to 3.6960)	2,909 (1 study[23])	⊕⊕⊖⊖ **low**[7,10,24,25,26]
LDCT screening vs. CXR screening (national death registry, death certificates, adjudicated cause[1]; follow-up: median 6.5 years)	**16,572**	**13,322** (11,600 to 15,301)	**RR 0.8039** (0.7000 to 0.9233)	53,454 (1 study[27])	⊕⊕⊕⊕ **high**[7,10,28,29,30]

*The assumed risk is the mean control group risk across studies. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

Footnotes for the GRADE Evidence Profile and Summary of Findings Tables for the Effect of Lung Cancer Screening on Lung Cancer Mortality

[1] In the Kaiser study[44] mortality was assessed using California state mortality records and codes from International Classification of Diseases, Adapted 8th revision. The PLCO trial[38] was linked to the National Death Index and used death certificates to confirm death and provisional cause (end point adjudication used to assign cause). The Mayo Lung Project[39] used the National Death Index-PLUS to assign vital status and date and cause of death. In the Czech study[42] 35% of deaths had autopsies which were used to confirm cause of death, otherwise cause was taken as given on death certificates. In the Memorial Sloan-Kettering[45] and Johns Hopkins[43] studies mortality was assessed using death certificates and death related clinical records. The North London study[46] established cause of death from hospital records and used the national death registry (General Register Office at Somerset House).

[2] The 2 RCTs are: the Kaiser Foundation study[44] and the PLCO study[38]

[3] Using Cochrane's Risk of Bias tool,[29] for this outcome, 1 study was rated as having low risk of bias[38] and the other was rated as having high risk of bias.[44] Both studies had low risk ratings for blinding of outcome assessment and selective reporting and both studies had unclear ratings for allocation concealment (not discussed in papers). The PLCO study had low risk ratings for random sequence generation, incomplete reporting of outcomes and other sources of bias, while the Kaiser study received a high risk of bias rating for random sequence generation (system used sequential case numbers), incomplete reporting (36%

attrition) and other sources of bias (control group contamination; baseline differences in outcome related risk factors). Given the unclear and high potential for bias, this body of evidence was downgraded for serious study limitations.

[4] The statistical heterogeneity is low [Chi^2=0.38, df=1 (P=0.54); I^2=0%], and although the direction of the effect is not consistent across studies, the confidence intervals overlap. This body of evidence was not downgraded for inconsistency.

[5] Both studies included mixed gender populations with about equal representation of men and women. The Kaiser study targeted middle-aged adults (35-54 years) and at outset the sample was about equally divided above and below age 45. The PLCO study targeted older adults (55-74 years) and enrolled about 33% of participants in their 50s, 53% in their 60s, and 13% in their early 70s. Both studies recruited smokers (current and former) as well as non-smokers; only 17% of the Kaiser study sample were identified as smokers (though study authors suggested this was likely an underestimate of actual smokers) while in the PLCO study 52% were identified as current or former smokers. The screening test in both studies was CXR. In the Kaiser study screening participants were offered CXR annually for 4 years while the PLCO participants were offered an annual CXR for 16 years as part of a comprehensive health check-up. The control group in both studies received usual care (screening was not offered or advised, but patients could be screened if they and/or their health care provider initiated testing). The length of follow-up for lung cancer mortality was up to 16 years in the Kaiser study and up to 13 years (median 11.9, mean 11.2, interquartile range 10-13) in the PLCO trial. Both studies were conducted in the US. The Kaiser study was initiated in 1964 while the PLCO trial started in 1993. There were no serious concerns regarding indirectness for this body of evidence.

[6] The sample size is adequate (82,583 CXR screening, 82,992 usual care), and the number of events is sufficient (1,257 CXR screening, 1,272 usual care) but the pooled effect estimate is not precise with a confidence interval that includes the no effect value [RR 0.9908 (95% CI 0.9171, 1.0705)]. This body of evidence was downgraded for serious concerns regarding imprecision.

[7] There were too few studies (n<10) to assess publication bias.[176]

[8] The single RCT is: the North London study[46]

[9] Using Cochrane's Risk of Bias tool,[29] for this outcome, this study was rated as having unclear risk of bias. Low risk ratings were assigned for random sequence generation, incomplete reporting, selective reporting and other risk of bias. Neither allocation concealment nor blinding of outcome assessment were addressed in the paper and thus these domains were assigned unclear ratings. Given the unclear potential for bias, this body of evidence was downgraded for serious study limitations.

[10] Single study therefore no inconsistency.

[11] This study included only men. The study targeted adults aged 40-70 years; about one-quarter of participants who were enrolled were aged 40-44, another quarter were aged 45-49, a bit less than a quarter (22%) were aged 50-54, and the remaining men were mostly in their late 50s (16%) or 60s (10%). At enrollment most of the sample (69%) were smokers, 19% were former smokers, and 12% were never smokers. The screening test was CXR. Men were enrolled only if they had negative results on a preliminary screen. Participants randomized to the more intense screening arm were then offered CXR every 6 months for a period of 3 years (up to 6 additional screens) while those assigned to the less intense screening arm had the eligibility CXR and then a second CXR 3 years later. The length of follow-up for lung cancer mortality was at least 5 years. This study was conducted in the UK during the 1960s. There were no serious concerns regarding indirectness for this body of evidence.

[12] The sample size is adequate (29,723 more intensive CXR screening, 25,311 less intensive CXR screening), but the number of events is insufficient (82 more intensive CXR screening, 68 less intensive CXR screening) and the pooled effect estimate is not precise with a confidence interval that includes the no effect value [RR 1.0269 (95% CI 0.7449, 1.4156)]. This body of evidence was downgraded for serious concerns regarding imprecision.

[13] The 4 RCTs are: the Czech Study,[42] the Johns Hopkins study,[43] the Mayo Lung Project[39] and the Memorial Sloan-Kettering study[45]

[14] Using Cochrane's Risk of Bias tool,[29] for this outcome, 2 studies were rated as having low risk of bias[43, 45] and the other 2 were rated as having unclear risk of bias.[39, 42] Three studies had low risk ratings for random sequence generation but 1 study did not clarify the approach for randomization. None of the studies discussed allocation concealment (unclear ratings) however for 1 study the Cochrane review[25] authors confirmed a low risk rating for this domain via communication with the researchers. Three studies stated that outcome assessment was blinded and had adequate attrition rates while the fourth study did not address either issue. All studies received low risk ratings for selective reporting and other risk of bias. Given the unclear potential for bias, this body of evidence was downgraded for serious study limitations.

[15] Although statistical heterogeneity is moderate [Chi^2=7.27, df=3 (P=0.06); I^2=59%], and the direction of the effect is not consistent across studies, the confidence intervals overlap. This body of evidence was not downgraded for inconsistency.

[16] All 4 studies included only men. At enrollment, about one-quarter of the Mayo Lung Project sample was under age 50, another quarter was aged 50-54, two-fifths were aged 55-65, and the remaining 10% were aged 65-70+. In the Czech study about 38% of the men were enrolled in their 40s, about half (48%) were in their 50s, and 14% were in their early 60s. About one-third of the Memorial Sloan-Kettering sample was under age 50 when recruited, a bit more than half (54%) of the men were in their 50s, and the remaining quarter were mostly in their 60s, with a small percentage (3%) in their 70s. In the Johns Hopkins study just under one-third (31%) of the men were in their late 40s, about a quarter (27%) were in their early 50s and another quarter (25%) were in their late 50s, and the remaining enrollees were mostly in their 60s. All 4 studies recruited only heavy smokers (current or recently quit). In the Mayo Lung study 94% of the men smoked at least 1 pack/day, 97% had smoked for 20+ years and 91% had 25+ pack-years. Similarly, all men in the Memorial Sloan-Kettering study smoked at least 1 pack/day and 91% had 25+ pack-years. Likewise all men in the Johns Hopkins study smoked at least 1 pack/day and 94% had 25+ pack-years. Half of the men in the Czech study had a lifetime consumption of 150,000 to 250,000 cigarettes (approximately 20.55 to 34.25 pack-years) and the other half had smoked more. All 4 studies used both CXR and SC as a dual screening test. Two studies only enrolled participants who had negative results on a preliminary screen. In the Mayo Lung Project participants were offered CXR and SC every 4 months for a period of 6 years (up to 18 screens); the less intense screening arm had the eligibility CXR and SC test and was subsequently advised to receive at least annual screening but these men were not offered systematic re-screening. The Czech study offered one group the dual screen every 6 months for 3 years (up to 6 screens) followed by 3 years of annual CXR testing; the less intense screening group received the dual test prior to randomization and again 3 years later after which annual CXR was offered for 3 more years. The Memorial Sloan-Kettering and Johns Hopkins studies offered half of the men annual CXR and SC every 4 months for 5 years (initial recruits may have had an additional 2-3 years of screening), and the other half were offered annual screening with CXR alone for the same amount of time. In terms of follow-up, lung cancer mortality was assessed at a median of 20.5 years in the Mayo Lung Project, for up to 15 years in the Czech study, and for up to 9 years in the Memorial Sloan-Kettering and Johns Hopkins studies. All 4 studies were initiated in the 1970s. Three studies were conducted in the US and 1 in Czechoslovakia. There were no serious concerns regarding indirectness for this body of evidence.

[17] The sample size is adequate (17,983 more intensive CXR+SC screening, 18,000 less intensive screening), and the number of events is sufficient (840 more intensive CXR+SC screening, 812 less intensive screening) but the pooled effect estimate is not precise with a confidence interval that includes the no effect value [RR 1.0141 (95% CI 0.8717, 1.1798)]. This body of evidence was downgraded for serious concerns regarding imprecision.

[18] The 3 RCTs are: the DANTE study,[21] the DLCST study[40] and the MILD study[23]

[19] Using Cochrane's Risk of Bias tool,[29] for this outcome, all 3 studies were rated as having unclear risk of bias Two studies had low risk ratings for random sequence generation and 1 was rated as unclear since the authors did not clarify the randomization approach. Allocation concealment was not discussed for any study therefore this domain was consistently rated as unclear. Blinding of mortality assessment was only confirmed in 1 study (low risk); in the other two studies blinding was not mentioned or was clearly not done. Low risk ratings were assigned across all 3 studies for incomplete and selective outcome reporting. For other sources of bias, 1 study was rated as unclear (control group contamination was not addressed) and the other 2 were assigned high risk ratings (in both cases for baseline differences in outcome related risk factors). Given the unclear and high potential for bias, this body of evidence was downgraded for serious study limitations.

[20] Although the statistical heterogeneity is moderate [Chi^2=2.96, df=2 (P=0.23); I^2=32%] and the direction of the effect is not consistent across studies, the confidence intervals overlap. This body of evidence was not downgraded for inconsistency.

[21] Two studies included mixed gender samples; slightly more than half of the participants in the DLCST study and about two-thirds of the MILD study sample were men. The DANTE study included only men. At enrollment, in the two mixed gender trials participants were in their late 50s (MILD median age 57 years; DLCST mean age 58 years); enrollees were slightly older (mean 64 years) in the male-only DANTE trial. All three studies recruited only current or former (quit in last 10 years) smokers. In the DANTE study a little over half of the men were current smokers and all the men had at least 20 pack-years (mean 47 pack-years). All participants in the DLCST study also had 20+ pack-years (mean 36 pack-years) and the mean number of cigarettes smoked per day was 19. Smoking history was a bit heavier in the MILD study with 90% of control and 69% of screening participants smoking upon enrollment, a mean of 38 pack-years across participants with consumption for two-thirds of the group at 20+ cigarettes per day. All three studies used annual LDCT as the screening test. In one intervention arm of the MILD study, participants were offered annual LDCT screening (median number of screens 5); the control group received usual care. Participants in the intervention arm of the DLCST trial were offered 5 annual LDCT screens; the control group received usual care. At baseline, men in the screening group of the DANTE trial had a CXR plus SC test in addition to their first of 5 annual LDCT scans; the usual care group also had the baseline CXR and SC test (no LDCT) but then only attended yearly for clinical reviews (no screening). In terms of follow-up, lung cancer mortality was assessed at a median of 4.4 years (maximum 6 years) in the MILD study, for 5 years in the DLCST study and for a median of 2.8 years (range 1.8 to 79.2 months) in the on-going DANTE trial. All 3 studies were initiated in the last 10-15 years. Two studies were conducted in Italy and 1 in Denmark. There were no serious concerns regarding indirectness for this body of evidence.

[22] The sample size is adequate (4,518 LDCT screening, 4,971 usual care), but the number of events is insufficient (47 LDCT screening, 38 usual care) and the pooled effect estimate is not precise with a confidence interval that includes the no effect value [RR 1.3460 (95% CI 0.7904, 2.2922)]. This body of evidence was downgraded for serious concerns regarding imprecision.

[23] The single RCT is: the MILD study[23]

[24] Using Cochrane's Risk of Bias tool,[29] for this outcome, this study was rated as having unclear risk of bias. Unclear ratings were applied to random sequence generation and allocation concealment, neither of which was explicitly addressed. Low risk ratings were assigned for blinding of outcome assessment, and

incomplete and selective outcome reporting. A high risk rating was applied for other sources of bias because of baseline differences in outcome related risk factors and lack of information on control group contamination. Given the unclear and high potential for bias, this body of evidence was downgraded for serious study limitations.

[25] The MILD trial included mixed gender participants; about two-thirds of the sample was men. At enrollment, participants were mostly in their late 50s (median age 57 years). Only current or former (quit in last 10 years) smokers were recruited; 90% of control and 69% of screening participants were current smokers, the overall mean pack-years was 38 and consumption for two-thirds of the sample was 20+ cigarettes per day. For this comparison, biennial LDCT was the screening test. In the intervention arm participants were offered LDCT screening every two years (median number of screens 3); the control group received usual care. In terms of follow-up, lung cancer mortality was assessed at a median of 4.4 years (maximum 6 years). This trial was initiated in 2005 in Italy. There were no serious concerns regarding indirectness for this body of evidence.

[26] The sample size is adequate (1,186 biennial LDCT screening, 1,723 usual care), but the number of events is insufficient (6 biennial LDCT screening, 7 usual care) and the pooled effect estimate is not precise with a confidence interval that includes the no effect value [RR 1.2452 (95% CI 0.4195, 3.6960)]. This body of evidence was downgraded for serious concerns regarding imprecision.

[27] The single RCT is: the NLST study[10]

[28] Using Cochrane's Risk of Bias tool,[29] for this outcome, this study was rated as having low risk of bias. Low risk ratings were applied to random sequence generation, blinding of outcome assessment, incomplete and selective outcome reporting and other potential sources. Only allocation concealment received an unclear rating because this was not explicitly addressed. Given the low potential for bias, this body of evidence was not downgraded for study limitations.

[29] The NLST study included mixed gender participants; about 60% of the sample was men. At enrollment, about 60% of participants were over age 60. Only current (48%) or former (quit in last 15 years) (52%) smokers were included. This trial compared two screening interventions: LDCT versus CXR. Participants in each group were offered 3 annual screens with their respectively assigned test. In terms of follow-up, lung cancer mortality was assessed at a median of 6.5 years (maximum 7.4 years). This trial was initiated in 2002 in the US. There were no serious concerns regarding indirectness for this body of evidence.

[30] The sample size is adequate (26,722 LDCT screening, 26,732 CXR screening), the number of events is sufficient (356 LDCT screening, 443 CXR) and the pooled effect estimate is precise with a narrow confidence interval [RR 0.8039 (95% CI 0.7000, 0.9233)]. This body of evidence was not downgraded for imprecision.

Forest Plot 1.1: Effect of Lung Cancer Screening Using CXR on Lung Cancer Mortality

Study or Subgroup	Screening Events	Total	Control Events	Total	Risk Ratio IV, Random, 95% CI	Risk Ratio IV, Random, 95% CI
1.1.1 CXR screening vs. usual care						
Kaiser Foundation	44	5,138	42	5,536	1.1288 [0.7408, 1.7198]	
PLCO Study	1,213	77,445	1,230	77,456	0.9863 [0.9117, 1.0671]	
Subtotal (95% CI)		**82,583**		**82,992**	**0.9908 [0.9171, 1.0705]**	
Total events	1,257		1,272			

Heterogeneity: Tau² = 0.00; Chi² = 0.38, df = 1 (P = 0.54); I² = 0%
Test for overall effect: Z = 0.23 (P = 0.82)

Study or Subgroup	Screening Events	Total	Control Events	Total	Risk Ratio IV, Random, 95% CI	Risk Ratio IV, Random, 95% CI
1.1.2 more intensive CXR screening vs. less intensive screening						
North London	82	29,723	68	25,311	1.0269 [0.7449, 1.4156]	
Subtotal (95% CI)		**29,723**		**25,311**	**1.0269 [0.7449, 1.4156]**	
Total events	82		68			

Heterogeneity: Not applicable
Test for overall effect: Z = 0.16 (P = 0.87)

Study or Subgroup	Screening Events	Total	Control Events	Total	Risk Ratio IV, Random, 95% CI	Risk Ratio IV, Random, 95% CI
1.1.3 more intensive CXR + SC screening vs. less intensive screening						
Czech Study	247	3,171	216	3,174	1.1446 [0.9600, 1.3646]	
Johns Hopkins	141	5,226	173	5,161	0.8049 [0.6466, 1.0020]	
Mayo Lung Project	337	4,618	303	4,593	1.1062 [0.9524, 1.2848]	
Memorial Sloan-Kettering	115	4,968	120	5,072	0.9784 [0.7599, 1.2598]	
Subtotal (95% CI)		**17,983**		**18,000**	**1.0141 [0.8717, 1.1798]**	
Total events	840		812			

Heterogeneity: Tau² = 0.01; Chi² = 7.27, df = 3 (P = 0.06); I² = 59%
Test for overall effect: Z = 0.18 (P = 0.86)

0.5 0.7 1 1.5 2
Favours Screening — Favours Control

Forest Plot 1.2: Effect of Lung Cancer Screening Using LDCT on Lung Cancer Mortality

Study or Subgroup	Screening Events	Screening Total	Control Events	Control Total	Risk Ratio IV, Random, 95% CI	Risk Ratio IV, Random, 95% CI
1.2.1 annual LDCT screening vs. usual care						
DANTE Study	20	1,276	20	1,196	0.9373 [0.5069, 1.7333]	
DLCST Study	15	2,052	11	2,052	1.3636 [0.6278, 2.9617]	
MILD Study-A	12	1,190	7	1,723	2.4821 [0.9801, 6.2860]	
Subtotal (95% CI)		**4,518**		**4,971**	**1.3460 [0.7904, 2.2922]**	
Total events	47		38			
Heterogeneity: Tau² = 0.07; Chi² = 2.96, df = 2 (P = 0.23); I² = 32%						
Test for overall effect: Z = 1.09 (P = 0.27)						
1.2.2 biennial LDCT screening vs. usual care						
MILD Study-B	6	1,186	7	1,723	1.2452 [0.4195, 3.6960]	
Subtotal (95% CI)		**1,186**		**1,723**	**1.2452 [0.4195, 3.6960]**	
Total events	6		7			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.40 (P = 0.69)						
1.2.3 LDCT screening vs. CXR screening						
NLST Study	356	26,722	443	26,732	0.8039 [0.7000, 0.9233]	
Subtotal (95% CI)		**26,722**		**26,732**	**0.8039 [0.7000, 0.9233]**	
Total events	356		443			
Heterogeneity: Not applicable						
Test for overall effect: Z = 3.09 (P = 0.002)						

0.1 0.2   0.5   1   2   5   10
Favours Screening   Favours Control

95

Evidence Set 2: KQ1 Benefits of Screening – All-Cause Mortality

- ES Table 2.1: GRADE Evidence Profile - Effect of Lung Cancer Screening on All-Cause Mortality

- ES Table 2.2: GRADE Summary of Findings - Effect of Lung Cancer Screening on All-Cause Mortality

- Forest Plot 2.1: Effect of Lung Cancer Screening Using CXR on All-Cause Mortality

- Forest Plot 2.2: Effect of Lung Cancer Screening Using LDCT on All-Cause Mortality

ES Table 2.1: GRADE Evidence Profile - Effect of Lung Cancer Screening on All-Cause Mortality*

No. of Studies	Design	Risk of Bias	Inconsistency	Indirectness	Imprecision	Other	No. of Participants		Effect				GRADE Quality Rating	Importance
							Screening	Control	Relative (95% CI)	Absolute per Million (Range)	ARR	NNS (95% CI)		
All-Cause Mortality: CXR screening vs. usual care (follow-up 10 to 16 years; assessed with: national or state death registries, adjudicated cause[1])														
2[2]	randomized trials	serious risk[3]	no serious inconsistency[4]	no serious indirectness[5]	serious imprecision[6]	none[7]	11,464/82,583 (13.8818%)	11,745/82,992 (14.1520%)	RR 0.9801 (0.9570 to 1.0037)	2,816 fewer (6,085 fewer to 524 more)	-	-	⊕⊕OO LOW	CRITICAL
All-Cause Mortality: more intensive CXR + SC screening vs. less intensive screening (follow-up 6 to 20.5 years; assessed with: national death registry, autopsies, death certificates, death-related clinical records[1])														
3[8]	randomized trials	serious risk[9]	no serious inconsistency[10]	no serious indirectness[11]	serious imprecision[12]	none[7]	3,327/12,757 (26.0798%)	3,229/12,839 (25.1499%)	RR 1.0391 (0.9727 to 1.1100)	9,834 more (6,866 fewer to 27,665 more)	-	-	⊕⊕OO LOW	CRITICAL
All-Cause Mortality: annual LDCT screening vs. usual care (follow-up 2.8 to 6 years; assessed with: national death registries, medical records[1])														
3[13]	randomized trials	serious risk[14]	no serious inconsistency[15]	no serious indirectness[16]	serious imprecision[17]	none[7]	138/4,518 (3.0544%)	107/4,971 (2.1525%)	RR 1.4161 (0.9051 to 2.2155)	8,956 more (2,043 fewer to 26,163 more)	-	-	⊕⊕OO LOW	CRITICAL
All-Cause Mortality: biennial LDCT screening vs. usual care (follow-up median 4.4 years; assessed with: national death registry[1])														
1[18]	randomized trial	serious risk[19]	no serious inconsistency[20]	no serious indirectness[21]	serious imprecision[22]	none[7]	20/1,186 (1.6863%)	20/1,723 (1.1608%)	RR 1.4528 (0.7851 to 2.6881)	5,256 more (2,494 fewer to 19,595 more)	-	-	⊕⊕OO LOW	CRITICAL
All-Cause Mortality: LDCT screening vs. CXR screening (follow-up median 6.5 years; assessed with: national death registry, death certificates, adjudicated cause[1])														
1[23]	randomized trial	no serious risk[24]	no serious inconsistency[20]	no serious indirectness[25]	no serious imprecision[26]	none[7]	1,877/26,722 (7.0242%)	2,000/26,732 (7.4817%)	RR 0.9389 (0.8836 to 0.9976)	4,571 fewer (180 fewer to 8,709 fewer)	0.46%	219 (115, 5,556)	⊕⊕⊕⊕ HIGH	CRITICAL

*Footnotes appear below the Summary of Findings Table

ES Table 2.2: GRADE Summary of Findings - Effect of Lung Cancer Screening on All-Cause Mortality

Outcome: All-Cause Mortality	Illustrative Comparative Risks* (95% CI)		Relative Effect (95% CI)	No. of Participants (Studies)	Quality of the Evidence (GRADE)
	Assumed Risk Number per Million Control	Corresponding Risk Number per Million Treatment			
CXR screening vs. usual care (national or state death registries, adjudicated cause[1]; follow-up: 10 to 16 years)	141,520	138,703 (135,434 to 142,043)	**RR 0.9801** (0.9570 to 1.0037)	165,575 (2 studies[2])	⊕⊕⊖⊖ **low**[3,4,5,6,7]
more intensive CXR + SC screening vs. less intensive screening (national death registry, autopsies, death certificates, death-related clinical records[1]; follow-up: 6 to 20.5 years)	251,499	261,333 (244,633 to 279,164)	**RR 1.0391** (0.9727 to 1.1100)	25,596 (3 studies[8])	⊕⊕⊖⊖ **low**[7,9,10,11,12]
annual LDCT screening vs. usual care (national death registries, medical records[1]; follow-up: 2.8 to 6 years)	21,525	30,481 (19,482 to 47,688)	**RR 1.4161** (0.9051 to 2.2155)	9,489 (3 studies[13])	⊕⊕⊖⊖ **low**[7,14,15,16,17]
biennial LDCT screening vs. usual care (national death registry[1]; follow-up: median 4.4 years)	11,608	16,864 (9,113 to 31,203)	**RR 1.4528** (0.7851 to 2.6881)	2,909 (1 study[18])	⊕⊕⊖⊖ **low**[7,19,20,21,22]
LDCT screening vs. CXR screening (national death registry, death certificates, adjudicated cause[1]; follow-up: median 6.5 years)	74,817	70,245 (66,108 to 74,637)	**RR 0.9389** (0.8836 to 0.9976)	53,454 (1 study[23])	⊕⊕⊕⊕ **high**[7,20,24,25,26]

*The assumed risk is the mean control group risk across studies. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

Footnotes for the GRADE Evidence Profile and Summary of Findings Tables for the Effect of Lung Cancer Screening on All-Cause Mortality

[1] The Kaiser study[44] used California state mortality records and codes from International Classification of Diseases, Adapted 8th revision. The PLCO trial[38] was linked to the National Death Index and used death certificates with end point adjudication to assign cause. The Mayo Lung Project[39] used the National Death Index-PLUS. In the Czech study[42] 35% of deaths had autopsies which were used to confirm cause of death, otherwise cause was taken as given on death certificates. The Memorial Sloan-Kettering[45] and Johns Hopkins[43] studies used death certificates and death related clinical records. The DANTE study[21] used hospital or family physician records. The DLCST study[40] used the Danish Civil Registration System and Danish causes of Death Register, as well as physician and hospital records, autopsies, police reports and the National Board of Health, with an adjudicating board to assign cause. The MILD study[23] used death certificates from the Instituto Nazionale de Statistical. The NLST study[10] used the National Death Index and death certificates with end point verification to assign cause.

[2] The 2 RCTs are: the Kaiser Foundation study[44] and the PLCO study[38]

[3] Using Cochrane's Risk of Bias tool,[29] for this outcome, 1 study was rated as having low risk of bias[38] and the other was rated as having high risk of bias.[44] Both studies had low risk ratings for blinding of outcome assessment and selective reporting and both studies had unclear ratings for allocation concealment (not discussed in papers). The PLCO study had low risk ratings for random sequence generation, incomplete reporting of outcomes and other sources of bias, while the Kaiser study received a high risk of bias rating for random sequence generation (system used sequential case numbers), incomplete reporting (36% attrition) and other sources of bias (control group contamination; baseline differences in outcome related risk factors). Given the unclear and high potential for bias, this body of evidence was downgraded for serious study limitations.

[4] The statistical heterogeneity is low [Chi^2=0.00, df=1 (P=1.00); I^2=0%], the direction of the effect is consistent across studies and the confidence intervals overlap. This body of evidence was not downgraded for inconsistency.

[5] Both studies included mixed gender populations with about equal representation of men and women. The Kaiser study targeted middle-aged adults (35-54 years) and at outset the sample was about equally divided above and below age 45. The PLCO study targeted older adults (55-74 years) and enrolled about 33% of participants in their 50s, 53% in their 60s, and 13% in their early 70s. Both studies recruited smokers (current and former) as well as never smokers; only 17% of the Kaiser study sample were identified as smokers (though study authors suggested this was likely an underestimate of actual smokers) while in the PLCO study 52% were identified as current or former smokers. The screening test in both studies was CXR. In the Kaiser study screening participants were offered CXR annually for 4 years while the PLCO participants were offered an annual CXR for 16 years as part of a comprehensive health check-up. The control group in both studies received usual care (screening was not offered or advised, but patients could be screened if they and/or their health care provider initiated testing). The length of follow-up for mortality outcomes was up to 16 years in the Kaiser study and up to 13 years (median 11.9, mean 11.2, interquartile range 10-13) in the PLCO trial. Both studies were conducted in the US. The Kaiser study was initiated in 1964 while the PLCO trial started in 1993. There were no serious concerns regarding indirectness for this body of evidence.

[6] The sample size is adequate (82,583 CXR screening, 82,992 usual care) and the number of events is sufficient (11,464 CXR screening, 11,745 usual care), but the pooled effect estimate is not precise with a confidence interval that includes the no effect value [RR 0.9801 (95% CI 0.9570, 1.0037)]. This body of evidence was downgraded for serious concerns regarding imprecision.

[7] There were too few studies (n<10) to assess publication bias.[176]

[8] The 3 RCTs are: the Czech study,[42] the Mayo Lung Project[39] and the Memorial Sloan-Kettering study[45]

[9] Using Cochrane's Risk of Bias tool,[29] for this outcome, 1 study was rated as having low risk of bias[45] and the other 2 were rated as having unclear risk of bias.[39, 42] Two studies had low risk ratings for random sequence generation but 1 study did not clarify the approach for randomization (unclear). None of the studies discussed allocation concealment (unclear ratings) however for 1 study the Cochrane review[25] authors confirmed a low risk rating for this domain via communication with the researchers. Two studies stated that outcome assessment was blinded and had adequate attrition rates while the third study did not address either issue. All studies received low risk ratings for selective reporting and other risk of bias. Given the unclear potential for bias, this body of evidence was downgraded for serious study limitations.

[10] Although the statistical heterogeneity is moderate [Chi^2=3.16, df=2 (P=0.21); I^2=37%], the direction of the effect is consistent across studies and the confidence intervals overlap. This body of evidence was not downgraded for inconsistency.

[11] All 3 studies included only men. At enrollment, about one-quarter of the Mayo Lung Project sample was under age 50, another quarter was aged 50-54, two-fifths were aged 55-65 and the remaining 10% were aged 65-70+. In the Czech study about 38% of the men were enrolled in their 40s, about half (48%) were in their 50s, and 14% were in their early 60s. About one-third of the Memorial Sloan-Kettering sample was under age 50 when recruited, a bit more than half (54%) of the men were in their 50s, and the remaining quarter were mostly in their 60s, with a small percentage (3%) in their 70s. All three studies recruited only heavy smokers (current or recently quit). In the Mayo Lung study 94% of the men smoked at least 1 pack/day, 97% had smoked for 20+ years and 91% had 25+ pack-years. Similarly, all men in the Memorial Sloan-Kettering study smoked at least 1 pack/day and 91% had 25+ pack-years. Half of the men in the Czech study had a lifetime consumption of 150,000 to 250,000 cigarettes (approximately 20.55 to 34.25 pack-years) and the other half had smoked more. All three studies used both CXR and SC as a dual screening test. Two studies only enrolled participants who had negative results on a preliminary screen. In the Mayo Lung Project participants were offered CXR and SC every 4 months for a period of 6 years (up to 18 screens); the less intense screening arm had the eligibility CXR and SC test and was subsequently advised to receive at least annual screening, but these men were not offered systematic re-screening. The Czech study offered the experimental group the dual screen every 6 months for 3 years (up to 6 screens) followed by 3 years of annual CXR testing; the group receiving less intense screening received the dual test prior to randomization and again 3 years later, after which annual CXR was offered for 3 more years. The Memorial Sloan-Kettering study offered half of the men annual CXR and SC every 4 months for 5-8 years, and the other half were offered annual screening with CXR alone for the same amount of time. In terms of follow-up, all-cause mortality was assessed at a median of 20.5 years in the Mayo Lung Project, for up to 9 years in the Memorial Sloan-Kettering study, and for 6 years in the Czech study. All 3 studies were initiated in the 1970s. Two studies were conducted in the US and 1 in Czechoslovakia. There were no serious concerns regarding indirectness for this body of evidence.

[12] The sample size is adequate (12,757 intensive CXR+SC, 12,839 less intense screening) and the number of events is sufficient (3,327 intensive CXR+SC, 3,229 less intense screening), but the pooled effect estimate is not precise with a confidence interval that includes the no effect value [RR 1.0391 (95% CI 0.9727, 1.1100)]. This body of evidence was downgraded for serious concerns regarding imprecision.

[13] The 3 RCTs are: the DANTE study,[21] the DLCST study[40] and the MILD study[23]

[14] Using Cochrane's Risk of Bias tool,[29] for this outcome, all 3 studies were rated as having unclear risk of bias. Two studies had low risk ratings for random sequence generation and 1 was rated as unclear since the authors did not clarify the randomization approach. Allocation concealment was not discussed for any study therefore this domain was consistently rated as unclear. Blinding of mortality assessment was only confirmed in 1 study (low risk); in the other two studies blinding was not mentioned or was clearly not done. Low risk ratings were assigned across all 3 studies for incomplete and selective outcome reporting. For other sources of bias, 1 study was rated as unclear (control group contamination was not addressed) and the other 2 were assigned high risk ratings (in both cases for baseline differences in outcome related risk factors). Given the unclear and high potential for bias, this body of evidence was downgraded for serious study limitations.

[15] The statistical heterogeneity is moderate [Chi^2=6.11, df=2 (P=0.05); I^2=67%], and although the direction of the effect is not consistent across studies, the confidence intervals overlap. This body of evidence was not downgraded for inconsistency.

[16] Two studies included mixed gender samples; slightly more than half of the participants in the DLCST study and about two-thirds of the MILD study sample were men. The DANTE study included only men. At enrollment, in the two mixed gender trials participants were in their late 50s (MILD median age 57 years; DLCST mean age 58 years); enrollees were slightly older (mean 64 years) in the male-only DANTE trial. All three studies recruited only current or former (quit in last 10 years) smokers. In the DANTE study a little over half of the men were current smokers and all the men had at least 20 pack-years (mean 47

pack-years). All participants in the DLCST study also had 20+ pack-years (mean 36 pack-years) and the mean number of cigarettes smoked per day was 19. Smoking history was a bit heavier in the MILD study with 90% of control and 69% of screening participants smoking upon enrollment, a mean of 38 pack-years across participants with consumption for two-thirds of the group at 20+ cigarettes per day. All three studies used annual LDCT as the screening test. In one intervention arm of the MILD study, participants were offered annual LDCT screening (median number of screens 5); the control group received usual care. Participants in the intervention arm of the DLCST trial were offered 5 annual LDCT screens; the control group received usual care. At baseline, men in the screening group of the DANTE trial had a CXR plus SC test in addition to their first of 5 annual LDCT scans; the usual care group also had the baseline CXR and SC test (no LDCT) but then only attended yearly for clinical reviews (no screening). In terms of follow-up, all-cause mortality was assessed at a median of 4.4 years (maximum 6 years) in the MILD study, for 5 years in the DLCST study and for a median of 2.8 years (range 1.8 to 79.2 months) in the on-going DANTE trial. All 3 studies were initiated in the last 10-15 years. Two studies were conducted in Italy and 1 in Denmark. There were no serious concerns regarding indirectness for this body of evidence.

[17] The sample size is adequate (4,518 LDCT screening, 4,971 usual care), but the number of events is insufficient (138 LDCT screening, 107 usual care) and the pooled effect estimate is not precise with a confidence interval that includes the no effect value [RR 1.4161 (95% CI 0.9051, 2.2155)]. This body of evidence was downgraded for serious concerns regarding imprecision.

[18] The single RCT is: the MILD study[23]

[19] Using Cochrane's Risk of Bias tool,[29] for this outcome, this study was rated as having unclear risk of bias. Unclear ratings were applied to random sequence generation and allocation concealment, neither of which was explicitly addressed. Low risk ratings were assigned for blinding of outcome assessment, and incomplete and selective outcome reporting. A high risk rating was applied for other sources of bias because of baseline differences in outcome related risk factors and lack of information on control group contamination. Given the unclear and high potential for bias, this body of evidence was downgraded for serious study limitations.

[20] Single study therefore no inconsistency.

[21] The MILD trial included mixed gender participants; about two-thirds of the sample was men. At enrollment, participants were mostly in their late 50s (median age 57 years). Only current or former (quit in last 10 years) smokers were recruited; 90% of control and 69% of screening participants were current smokers, the overall mean pack-years was 38 and consumption for two-thirds of the sample was 20+ cigarettes per day. For this comparison, biennial LDCT was the screening test. In the intervention arm participants were offered LDCT screening every two years (median number of screens 3); the control group received usual care. In terms of follow-up, all-cause mortality was assessed at a median of 4.4 years (maximum 6 years). This trial was initiated in 2005 in Italy. There were no serious concerns regarding indirectness for this body of evidence.

[22] The sample size is adequate (1,186 biennial LDCT screening, 1,723 usual care), but the number of events is insufficient (20 biennial LDCT screening, 20 usual care) and the pooled effect estimate is not precise with a confidence interval that includes the no effect value [RR 1.4528 (95% CI 0.7851, 2.6881)]. This body of evidence was downgraded for serious concerns regarding imprecision.

[23] The single RCT is: the NLST study[10]

[24] Using Cochrane's Risk of Bias tool,[29] for this outcome, this study was rated as having low risk of bias. Low risk ratings were applied to random sequence generation, blinding of outcome assessment, incomplete and selective outcome reporting and other potential sources. Only allocation concealment received an unclear rating because this was not explicitly addressed. Given the low potential for bias, this body of evidence was not downgraded for study limitations.

[25] The NLST study included mixed gender participants; about 60% of the sample was men. At enrollment, about 60% of participants were over age 60. Only current (48%) or former (quit in last 15 years) (52%) smokers were included. This trial compared two screening interventions: LDCT versus CXR. Participants in each group were offered 3 annual screens with their respectively assigned test. In terms of follow-up, all-cause mortality was assessed at a median of 6.5 years (maximum 7.4 years). This trial was initiated in 2002 in the US. There were no serious concerns regarding indirectness for this body of evidence.

[26] The sample size is adequate (26,722 LDCT screening, 26,732 CXR screening), the number of events is sufficient (1,877 LDCT screening, 2,000 CXR screening) and the pooled effect estimate is precise with a narrow confidence interval [RR 0.9389 (95% CI 0.8836, 0.9976)]. This body of evidence was not downgraded for imprecision.

Forest Plot 2.1: Effect of Lung Cancer Screening Using CXR on All-Cause Mortality

Study or Subgroup	Screening Events	Total	Control Events	Total	Risk Ratio IV, Random, 95% CI	Risk Ratio IV, Random, 95% CI
2.1.1 CXR screening vs. usual care						
Kaiser Foundation	585	5,138	643	5,536	0.9803 [0.8823, 1.0892]	
PLCO Study	10,879	77,445	11,102	77,456	0.9801 [0.9563, 1.0044]	
Subtotal (95% CI)		**82,583**		**82,992**	**0.9801 [0.9570, 1.0037]**	
Total events	11,464		11,745			
Heterogeneity: Tau² = 0.00; Chi² = 0.00, df = 1 (P = 1.00); I² = 0%						
Test for overall effect: Z = 1.65 (P = 0.10)						
2.1.2 more intensive CXR + SC screening vs. less intensive screening						
Czech Study	341	3,171	293	3,174	1.1649 [1.0045, 1.3510]	
Mayo Lung Project	2,493	4,618	2,445	4,593	1.0141 [0.9763, 1.0534]	
Memorial Sloan-Kettering	493	4,968	491	5,072	1.0251 [0.9104, 1.1543]	
Subtotal (95% CI)		**12,757**		**12,839**	**1.0391 [0.9727, 1.1100]**	
Total events	3,327		3,229			
Heterogeneity: Tau² = 0.00; Chi² = 3.16, df = 2 (P = 0.21); I² = 37%						
Test for overall effect: Z = 1.14 (P = 0.25)						

0.5   0.7   1   1.5   2
Favours Screening   Favours Control

Forest Plot 2.2: Effect of Lung Cancer Screening Using LDCT on All-Cause Mortality

Study or Subgroup	Screening Events	Total	Control Events	Total	Risk Ratio IV, Random, 95% CI	Risk Ratio IV, Random, 95% CI
2.2.1 annual LDCT screening vs. usual care						
DANTE Study	46	1,276	45	1,196	0.9581 [0.6401, 1.4341]	
DLCST Study	61	2,052	42	2,052	1.4524 [0.9851, 2.1413]	
MILD Study-A	31	1,190	20	1,723	2.2442 [1.2855, 3.9182]	
Subtotal (95% CI)		**4,518**		**4,971**	**1.4161 [0.9051, 2.2155]**	
Total events	138		107			
Heterogeneity: Tau² = 0.10; Chi² = 6.11, df = 2 (P = 0.05); I² = 67%						
Test for overall effect: Z = 1.52 (P = 0.13)						
2.2.2 biennial LDCT screening vs. usual care						
MILD Study-B	20	1,186	20	1,723	1.4528 [0.7851, 2.6881]	
Subtotal (95% CI)		**1,186**		**1,723**	**1.4528 [0.7851, 2.6881]**	
Total events	20		20			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.19 (P = 0.23)						
2.2.3 LDCT screening vs. CXR screening						
NLST Study	1,877	26,722	2,000	26,732	0.9389 [0.8836, 0.9976]	
Subtotal (95% CI)		**26,722**		**26,732**	**0.9389 [0.8836, 0.9976]**	
Total events	1,877		2,000			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.04 (P = 0.04)						

0.5 0.7 1 1.5 2
Favours Screening   Favours Control

Evidence Set 3: KQ1 Benefits of Screening – Stage at Diagnosis

- Forest Plot 3.1: Effect of Lung Cancer Screening Using CXR on Stage at Diagnosis (Early Stage I & II NSCLC)

- Forest Plot 3.2: Effect of Lung Cancer Screening Using LDCT on Stage at Diagnosis (Early Stage I & II NSCLC)

- Forest Plot 3.3: Effect of Lung Cancer Screening Using CXR on Stage at Diagnosis (Late Stage III & IV NSCLC)

- Forest Plot 3.4: Effect of Lung Cancer Screening Using LDCT on Stage at Diagnosis (Late Stage III & IV NSCLC)

Forest Plot 3.1: Effect of Lung Cancer Screening Using CXR on Stage at Diagnosis (Early Stage I & II NSCLC) (total number of cancers as denominator)

	Screening		Control		Risk Ratio	Risk Ratio
Study or Subgroup	**Events**	**Total**	**Events**	**Total**	**IV, Random, 95% CI**	**IV, Random, 95% CI**
3.1.1 CXR screening vs. usual care						
PLCO Study	574	1,447	479	1,374	1.1379 [1.0335, 1.2528]	
Subtotal (95% CI)		**1,447**		**1,374**	**1.1379 [1.0335, 1.2528]**	
Total events	574		479			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.63 (P = 0.009)						
3.1.2 more intensive CXR + SC screening vs. less intensive screening						
Czech Study	55	108	36	82	1.1600 [0.8535, 1.5765]	
Johns Hopkins	106	169	101	176	1.0930 [0.9199, 1.2986]	
Mayo Lung Project	99	206	51	160	1.5077 [1.1540, 1.9698]	
Memorial Sloan-Kettering	70	131	73	135	0.9882 [0.7907, 1.2350]	
Subtotal (95% CI)		**614**		**553**	**1.1548 [0.9774, 1.3645]**	
Total events	330		261			
Heterogeneity: Tau² = 0.01; Chi² = 6.02, df = 3 (P = 0.11); I² = 50%						
Test for overall effect: Z = 1.69 (P = 0.09)						

```
        0.5      0.7      1       1.5     2
     Favours Control      Favours Screening
```

Forest Plot 3.2: Effect of Lung Cancer Screening Using LDCT on Stage at Diagnosis (Early Stage I & II NSCLC) (total number of cancers as denominator)

Study or Subgroup	Screening Events	Total	Control Events	Total	Risk Ratio IV, Random, 95% CI	Risk Ratio IV, Random, 95% CI
3.2.1 annual LDCT screening vs. usual care						
DANTE Study	37	61	14	35	1.5164 [0.9637, 2.3860]	
DLCST Study	46	65	7	17	1.7187 [0.9534, 3.0982]	
Subtotal (95% CI)		**126**		**52**	**1.5887 [1.1092, 2.2755]**	
Total events	83		21			
Heterogeneity: Tau² = 0.00; Chi² = 0.11, df = 1 (P = 0.74); I² = 0%						
Test for overall effect: Z = 2.53 (P = 0.01)						
3.2.2 LDCT screening vs. CXR screening						
NLST Study	593	1,040	363	929	1.4593 [1.3256, 1.6064]	
Subtotal (95% CI)		**1,040**		**929**	**1.4593 [1.3256, 1.6064]**	
Total events	593		363			
Heterogeneity: Not applicable						
Test for overall effect: Z = 7.71 (P < 0.00001)						

0.5 0.7 1 1.5 2

Favours Control   Favours Screening

Forest Plot 3.3: Effect of Lung Cancer Screening Using CXR on Stage at Diagnosis (Late Stage III & IV NSCLC) (total number of cancers as denominator)

Study or Subgroup	Screening Events	Screening Total	Control Events	Control Total	Risk Ratio IV, Random, 95% CI	Risk Ratio IV, Random, 95% CI
3.3.1 CXR screening vs. usual care						
PLCO Study	873	1,447	895	1,374	0.9262 [0.8749, 0.9805]	
Subtotal (95% CI)		**1,447**		**1,374**	**0.9262 [0.8749, 0.9805]**	
Total events	873		895			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.64 (P = 0.008)						
3.3.2 more intensive CXR + SC screening vs. less intensive screening						
Czech Study	53	108	46	82	0.8748 [0.6670, 1.1474]	
Johns Hopkins	63	169	75	176	0.8748 [0.6745, 1.1346]	
Mayo Lung Project	107	206	109	160	0.7624 [0.6440, 0.9026]	
Memorial Sloan-Kettering	61	131	62	135	1.0139 [0.7825, 1.3138]	
Subtotal (95% CI)		**614**		**553**	**0.8515 [0.7527, 0.9632]**	
Total events	284		292			
Heterogeneity: Tau² = 0.00; Chi² = 3.46, df = 3 (P = 0.33); I² = 13%						
Test for overall effect: Z = 2.56 (P = 0.01)						

0.5 0.7 1 1.5 2
Favours Screening   Favours Control

Forest Plot 3.4: Effect of Lung Cancer Screening Using LDCT on Stage at Diagnosis (Late Stage III & IV NSCLC) (total number of cancers as denominator)

Study or Subgroup	Screening Events	Screening Total	Control Events	Control Total	Risk Ratio IV, Random, 95% CI	Risk Ratio IV, Random, 95% CI
3.4.1 annual LDCT screening vs. usual care						
DANTE Study	24	61	21	35	0.6557 [0.4340, 0.9907]	
DLCST Study	19	65	10	17	0.4969 [0.2870, 0.8603]	
Subtotal (95% CI)		**126**		**52**	**0.5933 [0.4266, 0.8250]**	
Total events	43		31			
Heterogeneity: Tau² = 0.00; Chi² = 0.63, df = 1 (P = 0.43); I² = 0%						
Test for overall effect: Z = 3.10 (P = 0.002)						
3.4.2 LDCT screening vs. CXR screening						
NLST Study	447	1,040	566	929	0.7055 [0.6467, 0.7695]	
Subtotal (95% CI)		**1,040**		**929**	**0.7055 [0.6467, 0.7695]**	
Total events	447		566			
Heterogeneity: Not applicable						
Test for overall effect: Z = 7.87 (P < 0.00001)						

0.5 0.7 1 1.5 2
Favours Screening   Favours Control

Evidence Set 4: KQ2 Harms of Screening or Invasive Follow-up Testing - Overdiagnosis

- ES Table 4.1: Findings Summary - Overdiagnosis

- ES Table 4.2: GRADE Rating - Overdiagnosis

ES Table 4.1: Findings Summary - Overdiagnosis

Author, Year	Study	Threshold / Cut-off Value	Age (years)	Screening Test	Overdiagnosis % (95% CI)
CXR + SC					
Yankelevitz, 2003 [66]	Mayo Lung Project	TVDT >400 and >300 days	≥45	CXR + SC every 4 months	2.27% (0.40, 11.81) for TVDT >400
					4.55% (1.26, 15.14) for TVDT >300
Yankelevitz, 2003 [66]	Memorial Sloan-Kettering study	TVDT >400 and >300 days	≥45	annual CXR + SC every 4 months	6.98% (2.40, 18.61) for TVDT >400
					16.28% (8.12, 29.97) for TVDT >300
LDCT					
Duffy, 2014 [65]	UKLS	lead time 5.5 years mean sojourn time 2.06 years	50 to 75	annual LDCT	10.99% (10.08, 11.98)
				biennial LDCT	10.99% (9.99, 12.07)
Patz, 2014 [67]	NLST	lead time ≥5 years	55 to 74	annual LDCT	18.50% (5.40, 30.60) based on screen detected cancers
					11.0% (3.20, 18.20) based on diagnosed cancers
Sone, 2007 [52]	rural area in Japan	tumour size 30 mm	40 to 74	annual LDCT	13.33% (6.26, 26.18)
Veronesi, 2007 [68]	COSMOS study	TVDT ≥400 days	≥50	annual LDCT	25.83% (18.84, 34.33)

TVDT = tumor volume doubling time

ES Table 4.2: GRADE Rating - Overdiagnosis

Study Design	No. Studies	Screening Test	Age (years, range)	Overdiagnosis (% range)	GRADE Rating*
Observational	2	CXR + SC	50 to 75	2.27% to 16.28%	LOW
Observational	4	LDCT	46 to 75	10.99% to 25.83%	LOW

* All bodies of evidence received low GRADE ratings due to the observational nature of the included studies, the variation observed across cut-point or threshold values, frequency of screening and length of follow-up.

Evidence Set 5: KQ2 Harms of Screening or Invasive Follow-up Testing – Death

- ES Table 5.1: GRADE Rating - Death from Invasive Follow-up Testing

- Forest Plot 5.1: Death from Invasive Follow-up Testing

ES Table 5.1: GRADE Rating - Death from Invasive Follow-up Testing

Study Design	No. Studies	Screening Test	Sample Size (Events/Total)*	Proportion per 1,000 (95% CI)	GRADE Rating**
Observational	4	CXR	23/778	28.60 (16.02, 41.17)	LOW
Observational	3	CXR + SC	21/333	47.67 (23.86, 71.49)	LOW
Observational	7	LDCT	20/1,502	11.18 (5.07, 17.28)	LOW

*Total is based on number of patients undergoing invasive follow-up procedures (e.g., video-assisted thoracoscopic surgery, fine-needle aspiration biopsy or fine-needle aspiration cytology, thoracotomy, mediastinoscopy, surgical resection) as a result of screening.

**Data are from observational studies, and the GRADE evidence ranking starts at low for observational studies. The types of invasive follow-up procedures and length of follow-up varied across studies with some papers only reporting data for patients with lung cancer.

Forest Plot 5.1: Death from Invasive Follow-up Testing

Author	Deaths	Total Number of Patients Undergoing Invasive Follow-up Testing		Proportion (95% CI)
CXR				
NLST Team, 2011	10	379		0.02639 (0.01439, 0.04788)
Melamed, 1984	5	144		0.03472 (0.01492, 0.07870)
Wilde, 1989 (A)	3	104		0.02885 (0.00986, 0.08140)
Wilde, 1989 (B)	5	125		0.04000 (0.01720, 0.09023)
Dominioni, 2012	0	26		0.00000 (0.00000, 0.12873)
Subtotal				0.02860 (0.01602, 0.04117)
CXR + SC				
Melamed, 1984	4	144		0.02778 (0.01085, 0.06924)
Fontana, 1990 (A)	7	94		0.07447 (0.03654, 0.14581)
Fontana, 1990 (B)	6	51		0.11765 (0.05505, 0.23381)
Kubik, 1990	4	44		0.09091 (0.03592, 0.21159)
Subtotal				0.04767 (0.02386, 0.07149)
LDCT				
NLST Team, 2011	16	1,075		0.01488 (0.00918, 0.02404)
Infante, 2009	2	96		0.02083 (0.00573, 0.07281)
Crestanello, 2004	1	55		0.01818 (0.00322, 0.09606)
Rzyman, 2013	0	104		0.00000 (0.00000, 0.03562)
Veronesi, 2008	0	108		0.00000 (0.00000, 0.03435)
Petersen, 2012	0	49		0.00000 (0.00000, 0.07270)
Diederich, 2004	1	15		0.06667 (0.01187, 0.29817)
Subtotal				0.01118 (0.00507, 0.01728)
Overall				0.01618 (0.01082, 0.02153)

-0.1   0   0.1   0.2   0.3

A=more intensive screening arm; B=less intensive screening arm

Evidence Set 6: KQ2 Harms of Screening or Invasive Follow-up Testing – Major Complications or Morbidity

- ES Table 6.1: GRADE Rating - Major Complications or Morbidity from Invasive Follow-up Testing

- Forest Plot 6.1: Major Complications or Morbidity from Invasive Follow-up Testing

ES Table 6.1: GRADE Rating - Major Complications or Morbidity* from Invasive Follow-up Testing

Study Design	No. Studies	Screening Test	Sample Size (Events/Total)**	Proportion per 1,000 (95% CI)	GRADE Rating***
Observational	1	CXR	24/379	63.32 (42.92, 92.49)	LOW
Observational	4	LDCT	92/1,336	43.29 (32.00, 54.58)	LOW

*Not all studies provided details on type of major complications or morbidity (i.e., requiring hospitalization or medical intervention).

**Total is based on the number of patients undergoing invasive follow-up procedures (e.g., video-assisted thoracoscopic surgery, fine-needle aspiration biopsy or fine-needle aspiration cytology, thoracotomy, mediastinoscopy, surgical resection) as a result of screening.

***Data are from observational studies, and the GRADE evidence ranking starts at low for observational studies. The types of invasive follow-up procedures and length of follow-up varied across studies with some papers only reporting data for patients with lung cancer.

Forest Plot 6.1: Major Complications or Morbidity from Invasive Follow-up Testing

Author	Events	Total Number of Patients Undergoing Invasive Follow-up Testing		Proportion (95% CI)
LDCT				
NLST Team, 2011	84	1,075		0.07814 (0.06356, 0.09573)
Veronesi, 2008	4	108		0.03704 (0.01450, 0.09138)
Rzyman, 2013	0	104		0.00000 (0.00000, 0.03562)
Petersen, 2012	4	49		0.08163 (0.03220, 0.19189)
Subtotal				0.04329 (0.03200, 0.05458)
CXR				
NLST Team, 2011	24	379		0.06332 (0.04292, 0.09249)
Subtotal				0.06332 (0.03854, 0.08811)
Overall				0.04673 (0.03646, 0.05700)

-0.1   0   0.1   0.2   0.3
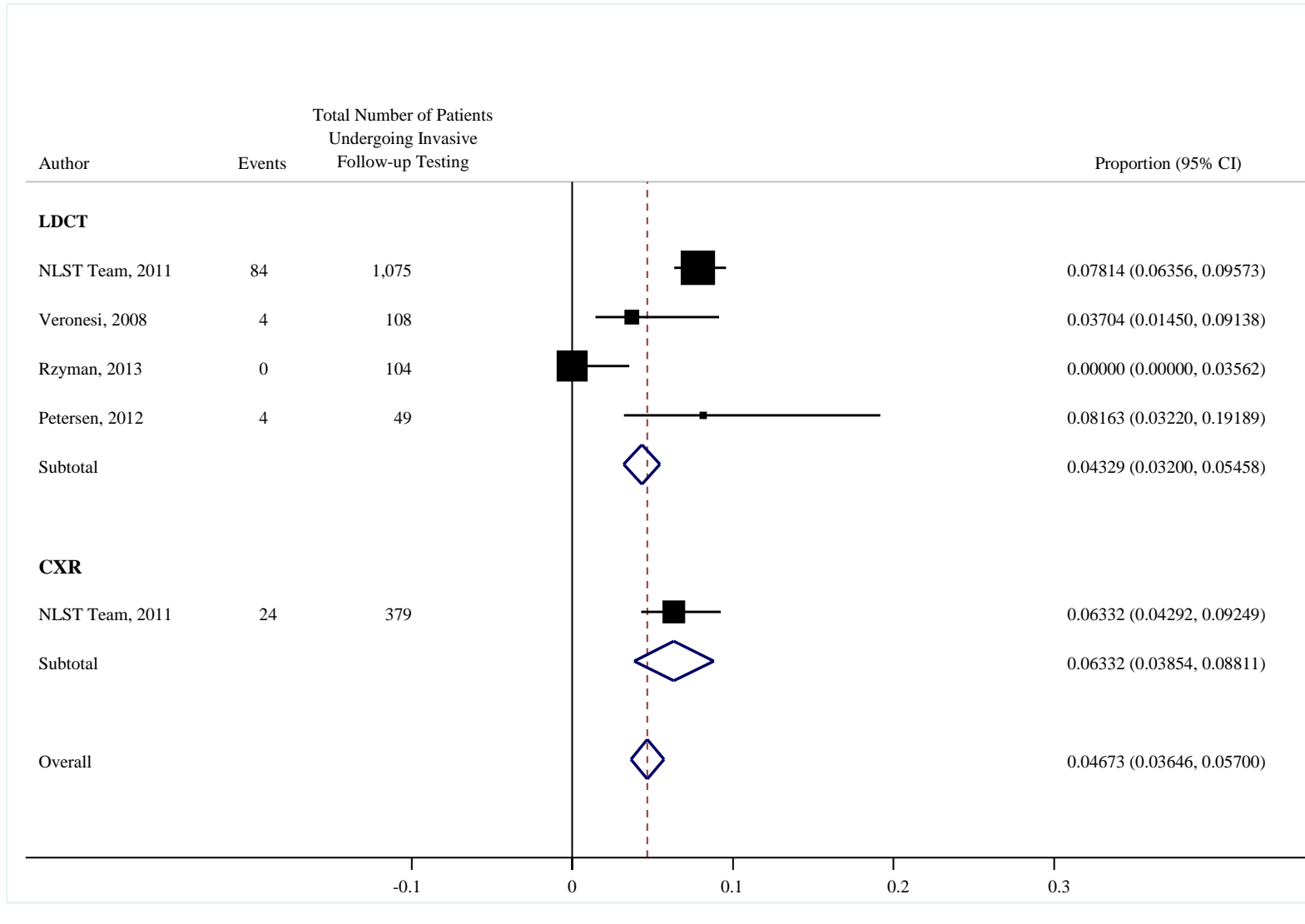
Evidence Set 7: KQ2 Harms of Screening or Invasive Follow-up Testing – False Positives

- ES Table 7.1: Findings Summary – False Positives

ES Table 7.1: Findings Summary – False Positives

Study Design	No. Studies	Screening Test	Screening Rounds	Sample Size (Events/Total*)	Proportion per 1,000 Median (Range)
Observational	3	CXR	multiple/repeat	2,098/33,199	65.0 (34.0 to 136.7)
Observational	2	LDCT	baseline	680/4,081	167.1 (79.0 to 255.3)
Observational	7	LDCT	multiple/repeat	8,290/42,774	233.0 (6.4 to 690.0)

* Total= number of people who underwent screening in these trials; Events=number of people who received at least one false positive result

Evidence Set 8: KQ2 Harms of Screening or Invasive Follow-up Testing – Consequences of False Positives

- ES Table 8.1: Findings Summary - Consequences of False Positives

- Forest Plot 8.1: Consequences of False Positives - Minor Invasive Procedures

- Forest Plot 8.2: Consequences of False Positives - Major Invasive Procedures

ES Table 8.1: Findings Summary – Consequences of False Positives

Study Design	No. Studies	Screening Test	Sample Size (Events/Total*)	Proportion per 1,000 (95% CI)
Minor Invasive Procedures**				
Observational	4	CXR	192/81,819	2.30 (1.49, 3.11)
Observational	7	LDCT	81/15,101	7.16 (3.27, 11.05)
Major Invasive Procedures***				
Observational	4	CXR	139/81,819	2.73 (0.96, 4.51)
Observational	17	LDCT	228/41,411	4.98 (3.68, 6.29)

* Total=number of people who underwent screening in these trials; Events=number of people with benign conditions subjected to minor/major invasive procedures as part of diagnostic follow-up

**Minor invasive procedures included fine-needle aspiration biopsy or fine-needle aspiration cytology, thoracic or lymph node biopsy, bronchoscopy, etc.

***Major invasive procedures included video-assisted thoracoscopic surgery, thoracotomy, surgical resection, etc.

Forest Plot 8.1: Consequences of False Positives - Minor Invasive Procedures

Author	Events (Screened Group)	Total (Screened Group)		Proportion (95% CI)
CXR				
Dominioni, 2012	1	1,244		0.00080 (0.00014, 0.00454)
Croswell, 2010	7	1,580		0.00443 (0.00215, 0.00912)
Hocking, 2013	179	77,445		0.00231 (0.00200, 0.00268)
Gohagan, 2004	5	1,550		0.00323 (0.00138, 0.00753)
Subtotal				0.00230 (0.00149, 0.00311)
LDCT				
Croswell, 2010	25	1,610		0.01553 (0.01054, 0.02282)
Henschke, 2001	4	1,000		0.00400 (0.00156, 0.01024)
Lopes Pegna, 2013	1	1,263		0.00079 (0.00014, 0.00447)
Horeweg, 2013	6	7,582		0.00079 (0.00036, 0.00173)
MacRedmond, 2006	5	449		0.01114 (0.00477, 0.02580)
Sobue, 2002	24	1,611		0.01490 (0.01003, 0.02207)
Gohagan, 2004	16	1,586		0.01009 (0.00622, 0.01632)
Subtotal				0.00716 (0.00327, 0.01105)
Overall				0.00381 (0.00233, 0.00528)

-0.01   0   0.01   0.02   0.03

Forest Plot 8.2: Consequences of False Positives - Major Invasive Procedures

Author	Events (Screened Group)	Total (Screened Group)		Proportion (95% CI)
CXR				
Dominioni, 2012	8	1,244		0.00643 (0.00326, 0.01264)
Croswell, 2010	4	1,580		0.00253 (0.00098, 0.00649)
Hocking, 2013	121	77,445		0.00156 (0.00131, 0.00187)
Gohagan, 2004	6	1,550		0.00387 (0.00178, 0.00842)
Subtotal				0.00273 (0.00096, 0.00451)
LDCT				
Croswell, 2010	8	1,610		0.00497 (0.00252, 0.00977)
Lopes Pegna, 2013	4	1,263		0.00317 (0.00123, 0.00811)
Horeweg, 2013	61	7,582		0.00805 (0.00627, 0.01032)
Pastorino, 2012	4	2,376		0.00168 (0.00065, 0.00432)
Rzyman, 2013	37	8,649		0.00428 (0.00311, 0.00589)
Saghir, 2011	7	2,052		0.00341 (0.00165, 0.00703)
Wilson, 2008	36	3,642		0.00988 (0.00715, 0.01365)
Infante, 2009	6	1,276		0.00470 (0.00216, 0.01022)
Pastorino, 2003	6	1,035		0.00580 (0.00266, 0.01259)
MacRedmond, 2006	1	449		0.00223 (0.00039, 0.01251)
Blanchon, 2007	3	336		0.00893 (0.00304, 0.02592)
Callol, 2007	2	406		0.00493 (0.00135, 0.01778)
Sobue, 2002	6	1,611		0.00372 (0.00171, 0.00810)
Veronesi, 2008	15	5,201		0.00288 (0.00175, 0.00475)
Crestanello, 2004	10	1,520		0.00658 (0.00358, 0.01207)
Diederich, 2004	4	817		0.00490 (0.00191, 0.01252)
Gohagan, 2004	18	1,586		0.01135 (0.00719, 0.01787)
Subtotal				0.00498 (0.00368, 0.00629)
Overall				0.00460 (0.00335, 0.00586)

-0.01 0 0.01 0.02 0.03

Evidence Set 9: KQ2 Harms of Screening or Invasive Follow-up Testing – Quality of Life

- Forest Plot 9.1: Health Related Quality of Life

Forest Plot 9.1: Health Related Quality of Life

Study	Screening			Control			Mean Difference	Mean Difference
	Mean	SD	Total	Mean	SD	Total	IV, Random, 95% CI	IV, Random, 95% CI
9.1.1 HRQOL: EQ-5D								
Mazzone, 2013	0.008	0.063	711	-0.028	0.063	713	0.0360 [0.0295, 0.0425]	
Subtotal (95% CI)			**711**			**713**	**0.0360 [0.0295, 0.0425]**	
Heterogeneity: Not applicable								
Test for overall effect: Z = 10.78 (P < 0.00001)								
9.1.2 HRQOL (Physical): SF-12								
PLCO study	-2.977	5.73	149	-2.7	5.3	179	-0.2770 [-1.4809, 0.9269]	
Subtotal (95% CI)			**149**			**179**	**-0.2770 [-1.4809, 0.9269]**	
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.45 (P = 0.65)								
9.1.3 HRQOL (Mental): SF-12								
PLCO study	-0.112	4.717	149	-1.3	4.299	179	1.1880 [0.2030, 2.1730]	
Subtotal (95% CI)			**149**			**179**	**1.1880 [0.2030, 2.1730]**	
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.36 (P = 0.02)								

```
        -4        -2         0         2         4
        Favours Screening    Favours Control
```

Evidence Set 10: KQ2 Harms of Screening or Invasive Follow-up Testing – Infections
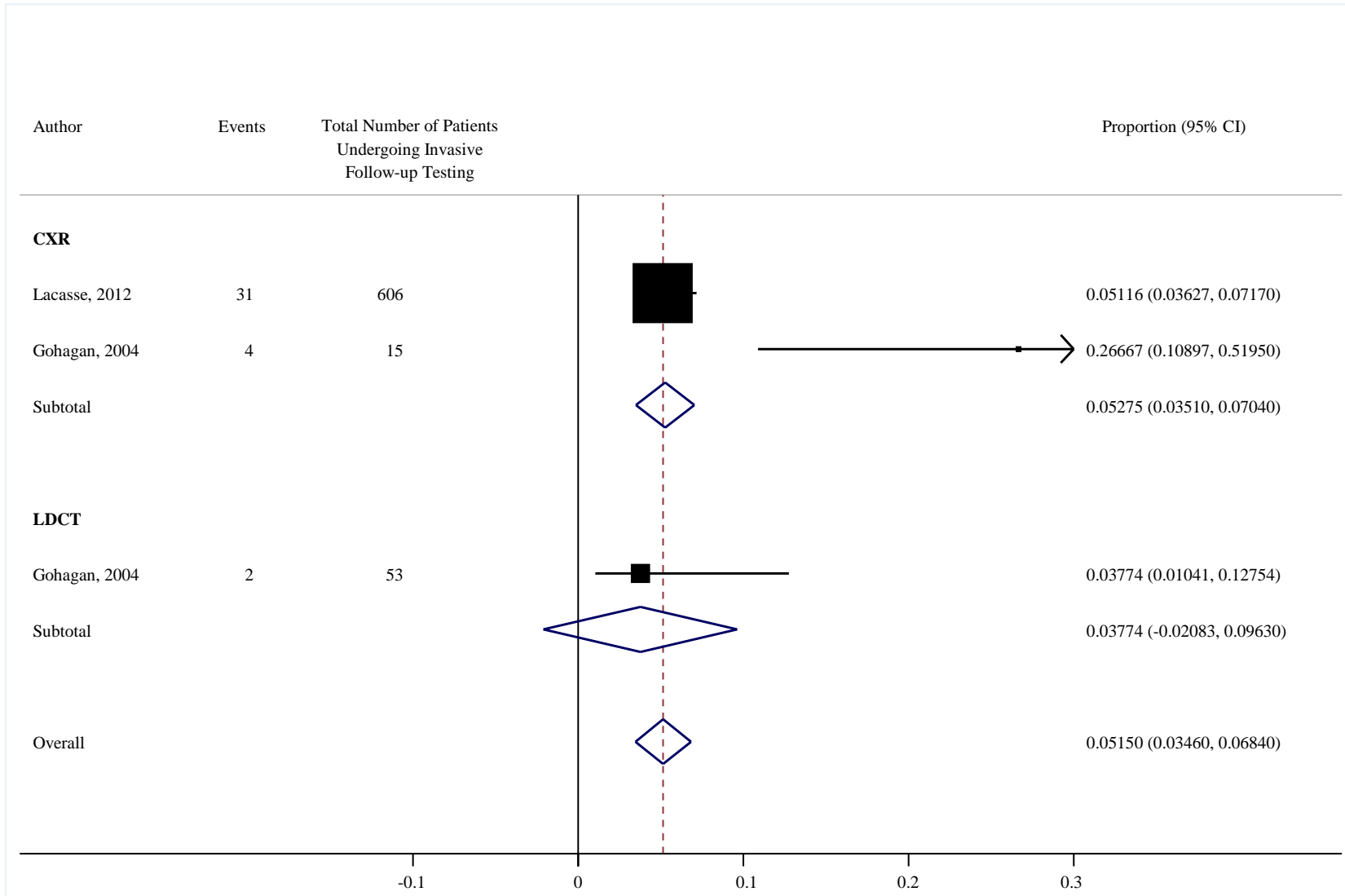
- ES Table 10.1: Findings Summary – Infections from Invasive Follow-up Testing

- Forest Plot 10.1: Infections from Invasive Follow-up Testing

ES Table 10.1: Findings Summary – Infections from Invasive Follow-up Testing

Study Design	No. Studies	Screening Test	Sample Size (Events/Total)*	Proportion per 1,000 (95% CI)
Observational	2	CXR	35/621	52.75 (35.10, 70.40)
Observational	1	LDCT	2/53	37.74 (10.41, 127.54)

*Total is based on the number of patients undergoing invasive follow-up testing (e.g., video-assisted thoracoscopic surgery, fine-needle aspiration biopsy or fine-needle aspiration cytology, thoracotomy, mediastinoscopy, surgical resection) as a result of screening.

Forest Plot 10.1: Infections from Invasive Follow-up Testing

Author	Events	Total Number of Patients Undergoing Invasive Follow-up Testing		Proportion (95% CI)
CXR				
Lacasse, 2012	31	606		0.05116 (0.03627, 0.07170)
Gohagan, 2004	4	15		0.26667 (0.10897, 0.51950)
Subtotal				0.05275 (0.03510, 0.07040)
LDCT				
Gohagan, 2004	2	53		0.03774 (0.01041, 0.12754)
Subtotal				0.03774 (-0.02083, 0.09630)
Overall				0.05150 (0.03460, 0.06840)

-0.1   0   0.1   0.2   0.3

Appendices

Appendix 1: Search Strategies for Key Questions and Contextual Questions

Appendix 2: Acknowledgements

Appendix 1: Search Strategies for Key Questions and Contextual Questions

Key Question 1 – Benefits of Screening

Database: MEDLINE In-Process & Other Non-Indexed Citations and MEDLINE: OVID
Last Searched: May 12, 2014
1. exp mass screening/
2. screen*.ti,ab.
3. 1 or 2
4. exp lung neoplasms/
5. lung neoplasm*.ti,ab.
6. lung cancer*.ti,ab.
7. lung.ti.
8. 4 or 5 or 6 or 7
9. 3 and 8
10. limit 9 to (english or french)
11. limit 10 to ed=20120523-20140512
12. (random* not non-random*).tw.
13. 11 and 12
14. limit 11 to (clinical trial or randomized controlled trial or controlled clinical trial)
15. 13 or 14

Database: EMBASE: OVID
Last Searched: May 13, 2014
1. exp screening/
2. screen*.ti,ab.
3. 1 or 2
4. exp lung cancer/
5. lung neoplasm*.ti,ab.
6. lung cancer*.ti,ab.
7. lung.ti.
8. 4 or 5 or 6 or 7
9. 3 and 8
10. limit 9 to (english or french)
11. limit 10 to human
12. limit 11 to em=201217-201417
13. limit 12 to (book or book series or conference abstract or conference paper or conference proceeding or "conference review" or editorial or letter or note)
14. 12 not 13
15. random?.tw. or placebo?.mp. or double-blind?.mp. or trial.mp. or control group.mp.
16. 14 and 15
17. limit 14 to (meta analysis or "systematic review")
18. limit 14 to (clinical trial or randomized controlled trial or controlled clinical trial)
19. 16 or 17 or 18

Database: Cochrane Central: OVID
Last Searched: May 13, 2014
1. exp mass screening/

2. screen*.ti,ab.
3. 1 or 2
4. exp lung neoplasms/
5. lung neoplasm*.ti,ab.
6. lung cancer*.ti,ab.
7. lung.ti.
8. 4 or 5 or 6 or 7
9. 3 and 8
10. limit 9 to yr="2012 - 2014"

Key Question 2 – Harms of Screening

Database: MEDLINE In-Process & Other Non-Indexed Citations and MEDLINE: OVID
Last Searched: May 13, 2014
1. exp mass screening/
2. screen*.ti,ab.
3. 1 or 2
4. exp lung neoplasms/
5. lung neoplasm*.ti,ab.
6. lung cancer*.ti,ab.
7. lung.ti.
8. 4 or 5 or 6 or 7
9. 3 and 8
10. Mass Screening/ae, ct, mo [Adverse Effects, Contraindications, Mortality]
11. mass chest x-ray/ae, ct, mo
12. 10 or 11
13. 8 and 12
14. ((adverse or undesirable or harm* or serious) adj3 (effect? or reaction? or event? or outcome?)).tw.
15. harm?.tw.
16. (overdiagnosis or over diagnosis or over-diagnosis or over detection or overdetection or over-detection or overtreatment or over treatment or over-treatment).tw.
17. Mortality/
18. adverse.tw.
19. (unnecessary adj3 treatment?).tw.
20. 14 or 15 or 16 or 17 or 18 or 19
21. 9 and 20
22. 13 or 21
23. limit 22 to (english or french)
24. limit 23 to yr="2000 -Current"
25. animals/ not (animals/ and humans/)
26. 24 not 25

Database: EMBASE: OVID
Last Searched: May 13, 2014
1. exp screening/
2. screen*.ti,ab.

3. 1 or 2
4. exp lung cancer/
5. lung neoplasm*.ti,ab.
6. lung cancer*.ti,ab.
7. lung.ti.
8. 4 or 5 or 6 or 7
9. 3 and 8
10. Mass Screening/ae
11. 8 and 10
12. ((adverse or undesirable or harm* or serious) adj3 (effect? or reaction? or event? or outcome?)).tw.
13. harm?.tw.
14. (overdiagnosis or over diagnosis or over-diagnosis or over detection or overdetection or over-detection or overtreatment or over treatment or over-treatment).tw.
15. adverse.tw.
16. (unnecessary adj3 treatment?).tw.
17. 12 or 13 or 14 or 15 or 16
18. 9 and 17
19. 11 or 18
20. limit 19 to (book or book series or conference abstract or conference paper or conference proceeding or "conference review" or editorial or letter or note)
21. 19 not 20
22. limit 21 to human
23. limit 22 to (english or french)
24. limit 23 to yr="2000 -Current"

Contextual Questions

Database: MEDLINE In-Process & Other Non-Indexed Citations and MEDLINE: OVID
Last Searched: June 12, 2014
1. exp mass screening/
2. screen*.ti,ab.
3. 1 or 2
4. exp lung neoplasms/
5. lung neoplasm*.ti,ab.
6. lung cancer*.ti,ab.
7. lung.ti.
8. 4 or 5 or 6 or 7
9. 3 and 8
10. exp continental population groups/
11. exp Ethnic Groups/
12. indians, north american/ or inuits/
13. first nations.tw.
14. (aboriginal? and canada).tw.
15. native canadians.tw.
16. (immigran* or new canadians).tw.
17. ((African or Asian or Indo or Columbian or Spanish or Chinese) adj2 Canadian?).mp.

18. Rural Population/
19. (rural adj (population? or area? or region?)).tw.
20. Rural Health/ or Rural Health Services/
21. Healthcare Disparities/
22. Social Class/
23. poverty/
24. socioeconomic.tw.
25. Socioeconomic Factors/
26. (poor or disadvantaged or poverty or social status).tw.
27. exp homeless persons/ or vulnerable populations/
28. exp "Costs and Cost Analysis"/
29. (cost or costs).tw.
30. *"patient acceptance of health care"/ or *patient compliance/ or *patient participation/ or patient satisfaction/ or patient preference/ or *treatment refusal/
31. ((parent? or guardian*) adj3 (acceptance or preference? or satisfaction or experience?)).tw.
32. (consumer? adj3 (acceptance or preference? or satisfaction or experience?)).tw.
33. (patient? adj3 (acceptance or perference? or satisfaction or experience?)).tw.
34. willingness to pay.tw.
35. ((conjoint or contingent) adj3 (valuation or analysis)).tw.
36. exp Canada/
37. (Canada or Canadian or Ontario or British Columbia or Alberta or Saskatchewan or Manitoba or Quebec or Nova Scotia or Prince Edward Island or Newfoundland or New Brunswick or Yukon or Northwest Territories or Nunavut).tw.
38. (sensitivity or specificity or predictive values or likelihood ratios).mp.
39. exp "Sensitivity and Specificity"/
40. (false negative or false positive).mp.
41. (screen* adj3 (interval? or frequency)).tw.
42. or/10-41
43. 9 and 42
44. (Canada or Canadian or Ontario or British Columbia or Alberta or Saskatchewan or Manitoba or Quebec or Nova Scotia or Prince Edward Island or Newfoundland or New Brunswick or Yukon or Northwest Territories or Nunavut).ti.
45. 8 and 44
46. 43 or 45
47. limit 46 to (english or french)
48. limit 47 to yr="2009 -Current"

Database: EMBASE: OVID
Last Searched: June 12, 2014
1. (sensitivity or specificity or predictive values or likelihood ratios).mp.
2. exp "Sensitivity and Specificity"/
3. (false negative or false positive).mp.
4. (screen* adj3 (interval? or frequency)).tw.
5. exp "ethnic and racial groups"/
6. first nations.tw.
7. (aboriginal? and canada).tw.
8. native canadians.tw.

9. (immigran* or new canadians).tw.
10. ((African or Asian or Indo or Columbian or Spanish or Chinese) adj2 Canadian).mp.
11. rural health care/
12. rural population/
13. (rural adj (population? or area? or region?)).tw.
14. exp economic evaluation/
15. cost.tw.
16. exp patient attitude/
17. (women? adj3 (acceptance or preference? or satisfaction or experience?)).tw.
18. (consumer? adj3 (acceptance or preference? or satisfaction or experience?)).tw.
19. (patient? adj3 (acceptance or preference? or satisfaction or experience?)).tw.
20. willingness to pay.tw.
21. ((conjoint or contingent) adj3 (valuation or analysis)).tw.
22. or/16-21
23. exp socioeconomics/
24. exp social status/
25. (poor or disadvantaged or poverty or social status).tw.
26. health care disparity/
27. miscellaneous named groups/ or lowest income group/ or medically underserved/ or vulnerable population/
28. exp Canada/
29. (Canada or Canadian or Ontario or British Columbia or Alberta or Saskatchewan or Manitoba or Quebec or Nova Scotia or Prince Edward Island or Newfoundland or New Brunswick or Yukon or Northwest Territories or Nunavut).tw.
30. (Canada or Canadian or Ontario or British Columbia or Alberta or Saskatchewan or Manitoba or Quebec or Nova Scotia or Prince Edward Island or Newfoundland or New Brunswick or Yukon or Northwest Territories or Nunavut).ti.
31. exp screening/
32. screen*.ti,ab.
33. 31 or 32
34. exp lung cancer/
35. lung neoplasm*.ti,ab.
36. lung cancer*.ti,ab.
37. lung.ti.
38. 34 or 35 or 36 or 37
39. 33 and 38
40. or/1-29
41. 39 and 40
42. 30 and 38
43. 41 or 42
44. limit 43 to (english or french)
45. limit 44 to yr="2009 -Current"
46. limit 45 to (book or book series or conference abstract or editorial or letter or note)
47. 45 not 46
48. screening test/
49. 38 and 48

50. limit 49 to (english or french)
51. limit 50 to yr="2009 -Current"
52. limit 51 to (book or book series or conference abstract or editorial or letter or note)
53. 51 not 52
54. 47 or 53

Appendix 2: Acknowledgements