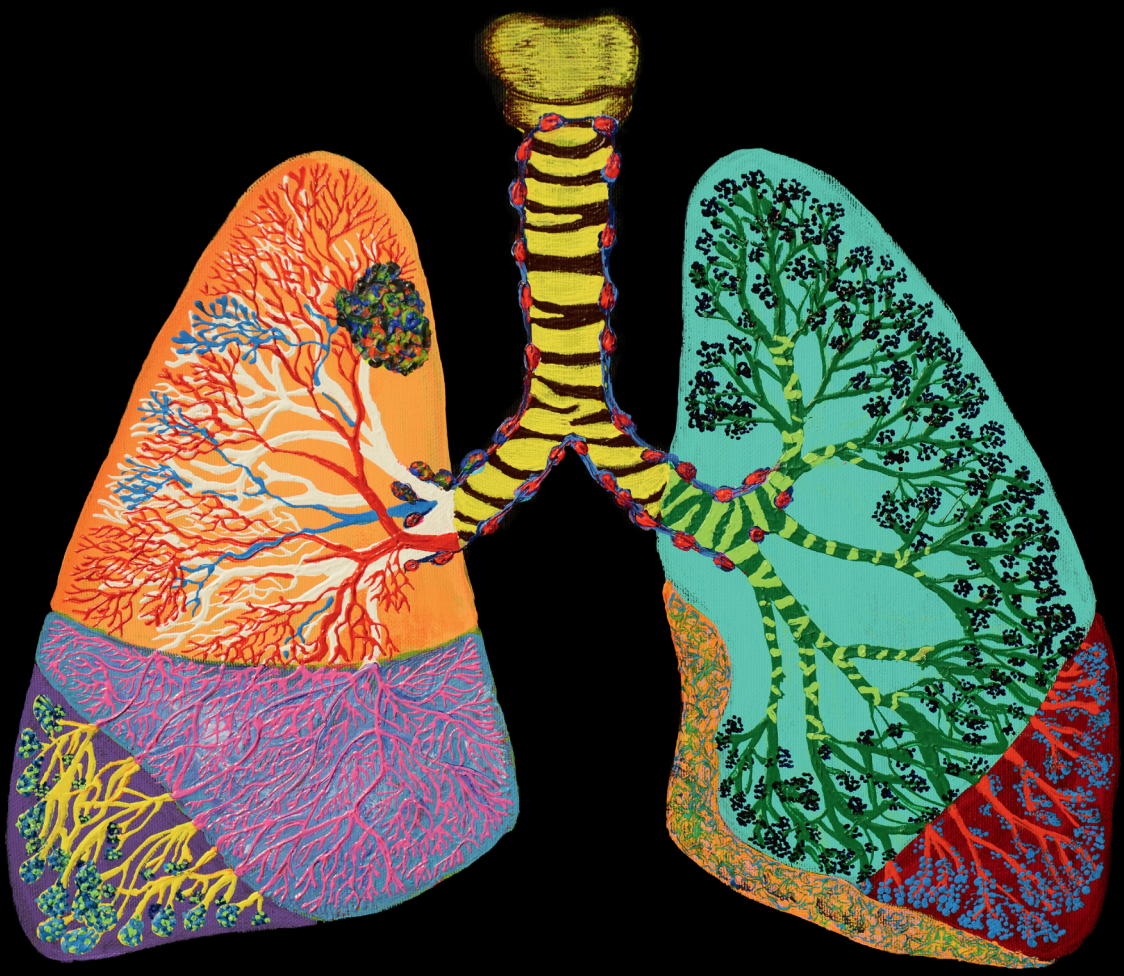
The NELSON Lung Cancer Screening Trial:
Final screening round and follow-up



A.U. Yousaf-Khan

The NELSON Lung Cancer Screening Trial:

Final screening round and follow-up

A.U. Yousaf-Khan

The NELSON Lung Cancer Screening Trial: Final screening round and follow-up

ISBN: 978-94-6332-365-9

Thesis, Erasmus University Rotterdam, Netherlands

Cover illustration: Anum & Shahryar (Instagram: sci.art_by_anum)
Lay-out and print: GVO drukkers en vormgevers B.V., Ede

This thesis was financially supported by the Department of Public Health (Erasmus MC) and Erasmus MC, Rotterdam.

The NELSON Lung Cancer Screening Trial:
Final screening round and follow-up

De NELSON longkankerscreening trial:
Laatste screeningsronde en follow-up

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

Prof.dr. H.A.P. Pols

en volgens besluit van het College voor Promoties.
De openbare verdediging zal plaatsvinden op

woensdag 27 juni 2018 om 13.30 uur

door

Assma Uraujh Yousaf-Khan
geboren te **Rotterdam**

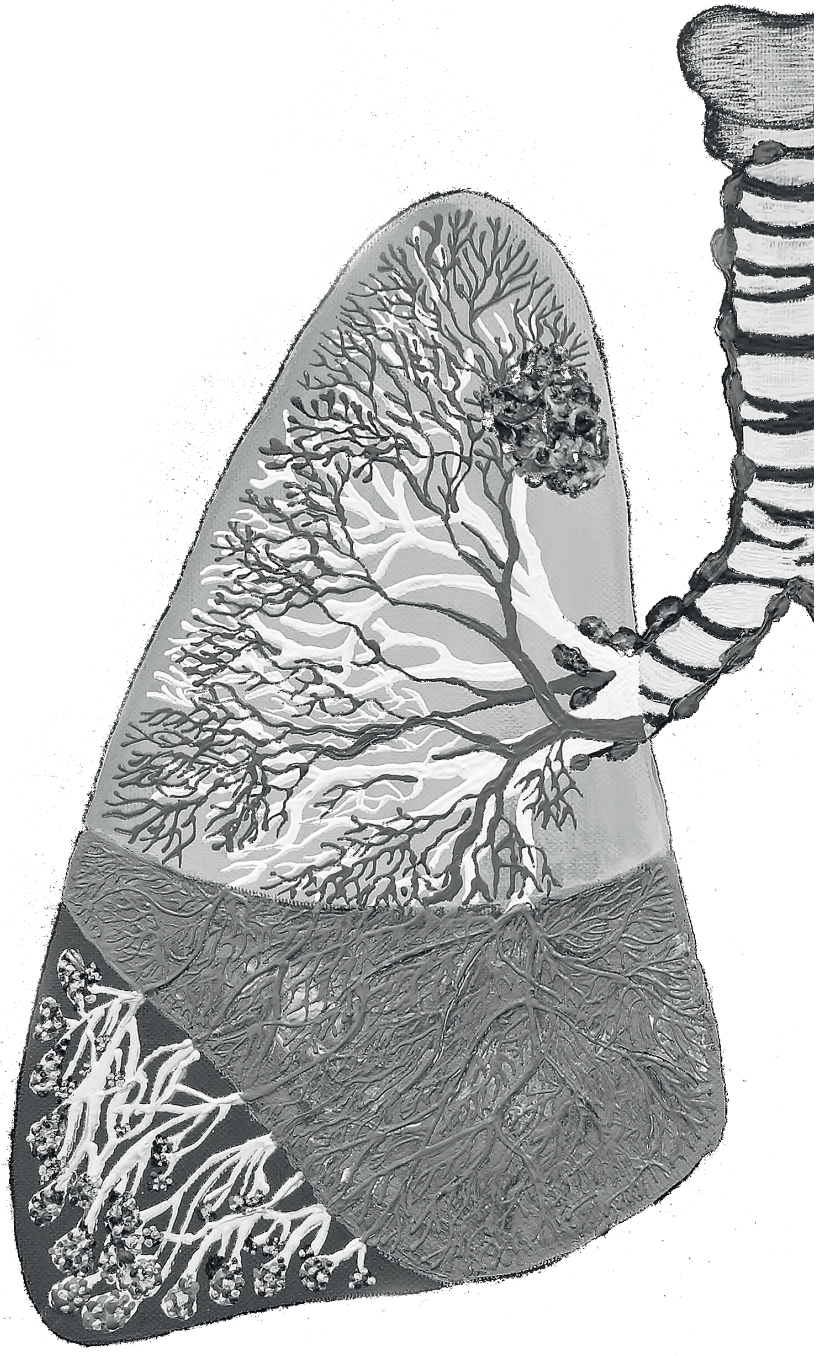
Erasmus University Rotterdam

Promotiecommissie

Promotor:	Prof. dr. H.J. de Koning
Overige leden:	Prof. dr. J.G.J.V. Aerts Prof. dr. P.E. Postmus Prof. dr. J.W.P.F. Kardaun
Copromotor:	dr. C.M. van der Aalst

CONTENTS

Chapter I	Introduction.	7
Part I: the optimization of the NELSON screening rounds		**41**
Chapter II	Final screening round of the NELSON lung cancer screening trial: the effect of a 2.5-year screening interval. *Thorax*	43
Chapter III	Risk stratification based on screening history: the NELSON lung cancer screening study. *Thorax*	69
Part II: interim stage shift results in the NELSON trial		**95**
Chapter IV	Cancer stage shift and treatment shift in the NELSON lung cancer screening trial: implications for clinicians. *Submitted*	97
Part III: the cause of death of the NELSON study participants		**115**
Chapter V	Baseline characteristics and mortality outcomes of NELSON control group participants and eligible nonresponders. *Journal of Thoracic Oncology*	117
Chapter VI	Blinded and uniform cause of death verification in a lung cancer CT screening trial *Lung Cancer*	135
Chapter VII	General discussion	165
	Summary	193
	Samenvatting	197
	Dankwoord	201
	About the author	205
	PhD Portfolio	207
	List of publications	211



CHAPTER I

Introduction

Epidemiology

Lung cancer is by far the leading cause of cancer death among males and the second leading cause of cancer death among females worldwide [1, 2]. Each year, more persons die from lung cancer than of colon, breast, and prostate cancers combined [2]. Most lung cancer cases are diagnosed in symptomatic patients. At that moment, the majority of patients have already reached an advanced stage of lung cancer, which is correlated with a poor survival [3].

In the Netherlands, lung cancer is the leading cause of death among both genders: approximately 25% of cancer deaths among males and 20% of cancer deaths among females are caused by lung cancer (**Figure 1**) [4].

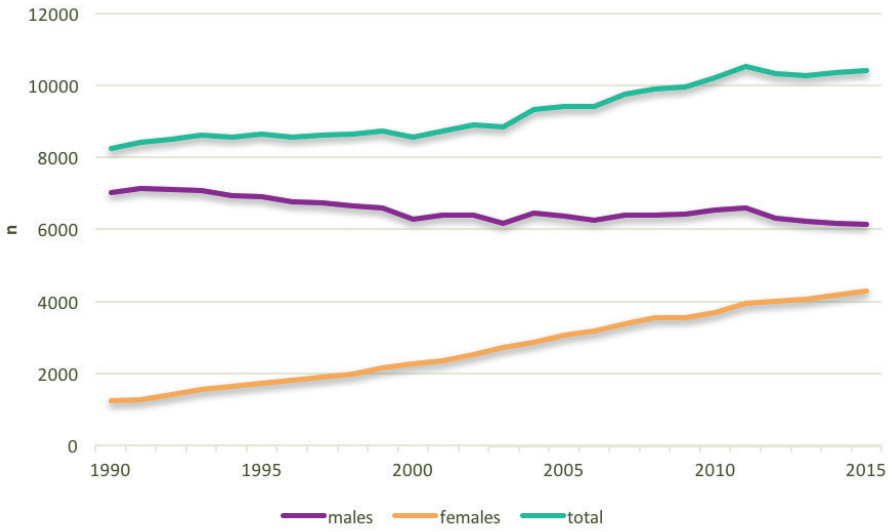


Figure 1: lung cancer deaths in the Netherlands from 1990 to 2015 [4]

Lung cancer incidence rates vary demographically by sex and age, and geographically by differences in historical smoking behaviours [1, 5]. Long term cigarette smoking accounts for approximately 85% of lung cancer cases [6-8]. Time lag between the onset of smoking and the development for lung cancer is approximately 20 to 30 years [9]. With the decrease of smoking prevalence in males since the 1960s, the prevalence of lung cancer among males has been decreasing worldwide [2]. Among females, the decline of smoking prevalence started in the late 1980s and plateaued in the 1990s. Hence, the prevalence of lung cancer among females is increasing [2].

In the Netherlands, 12,000 people are diagnosed with lung cancer every year, of whom approximately 60% are male (**Figure 2**) [10]. In the Netherlands, approximately 30% of males

and 22.6% of females do smoke, which corresponds to 26.3% of the Dutch population over 18 years of age [11, 12]. In 2015, the proportion of heavy smokers (people who smoke 20 or more cigarettes per day) decreased since 2000 from 35.0% to 15.3%.

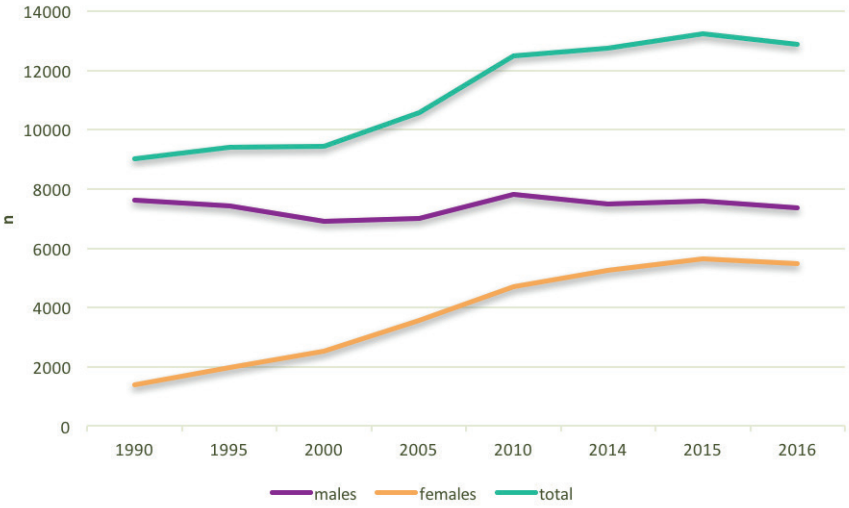


Figure 2: lung cancer incidence in the Netherlands, by gender [10]

Aetiology

In approximately 85% of all lung cancer cases, cigarette smoking is the primary cause for developing lung cancer [6-8, 13]. The amount and duration of cigarette smoking is strongly related to the development of lung cancer [6-8, 13]. In which, the duration of smoking is considered to be the strongest link with lung cancer risk [14]. Unadjusted lifetime risk for developing lung cancer is approximately 6.4% [3]. A current smoker who smoked one pack per day for 40 years has a 20 to 50 times higher risk to develop lung cancer than a never smoker [15]. In general, one out of nine smokers develops lung cancer [16]. Most cases of lung cancers occur in moderate to heavy smokers [16].

Other risk factors associated with lung cancer are: 1) family history of lung cancer, in which genetic factors can affect the risk for developing lung cancer regardless of exposure to cigarette smoking [17, 18]; 2) other lung diseases such as COPD, chronic bronchitis or emphysema which may increase the risk for developing lung cancer independently from smoking [19-21]; 3) passive smoking [22, 23]; 4) asbestosis and radon exposure [24-26]; and 5) a medical history of treatment with radiotherapy for a non-Hodgkin lymphoma or for breast cancer with increased risk for a second primary lung cancer [27, 28]

Histology

Lung cancer is defined as a malignant neoplasm of an unspecified part of an unspecified bronchus or lung according to the International Statistical Classification of Diseases and Related Health Problems (ICD-10; C34) [29]. Lung cancer usually arises from uncontrolled cell growth of the epithelium that lines in the bronchial tree and may spread to a site distant from the lungs and produce metastatic tumours in other parts of the body (e.g. brain, bones, liver or adrenal glands) [29].

There are two main types of lung cancer: small cell lung carcinoma (SCLC) and non-small cell lung carcinoma (NSCLC), accounting for approximately 15% and 85% of lung cancer cases, respectively [1]. These types are diagnosed based on the microscopic appearance of the malignant cells. The distinction between SCLC and NSCLC is essential for staging, treatment and prognosis of the lung cancer [30].

Small cell lung carcinoma

Approximately 15% of all lung cancers are SCLCs [29, 31]. They are characterised by small "blue" malignant cells about twice the size of lymphocytes (**Figure 3**). The cytoplasm is sparse and nuclear features include finely dispersed chromatin without distinct nucleoli. SCLC is histologically divided into two subtypes: oat cell carcinoma and a combined small cell carcinoma (usually a combination of SCLC with adenocarcinoma, squamous cell carcinoma or large cell carcinoma) [29, 30, 32]. Large cell neuroendocrine carcinoma (LCNEC) is officially classified under NSCLC, but its biological behaviour is similar to that of SCLC [32].

For SCLC and LCNEC there are standard immunohistochemical markers for lung origin and/or neuroendocrine features which are useful for establishing the diagnosis [33]. A majority of the SCLC express the thyroid transcription factor (TTF-1), which can help in distinguishing LCNEC from other neuroendocrine carcinomas [34]. Other markers that can be used to differentiate include CD56, chromogranin and synaptophysin [35]. Up to two-thirds of SCLC will be negative for chromogranin and synaptophysin and CD56 will be positive in approximately 90–100% of cases [35]. SCLC can also produce different kinds of hormones (e.g. ACTH and vasopressin) and antibodies which can cause various paraneoplastic neurological and endocrinologic syndromes [36, 37].

Lung cancer is a result of many genetic changes and exposure to tobacco and other possible carcinogens (e.g. radon, coal). In the case of SCLC, p53 mutations, loss of retinoblastoma gene (Rb) and strong expression of cKit are observed [33, 34]. Unlike NSCLC, mutations in the EGFR and KRAS oncogenes and p16 abnormalities are rare.

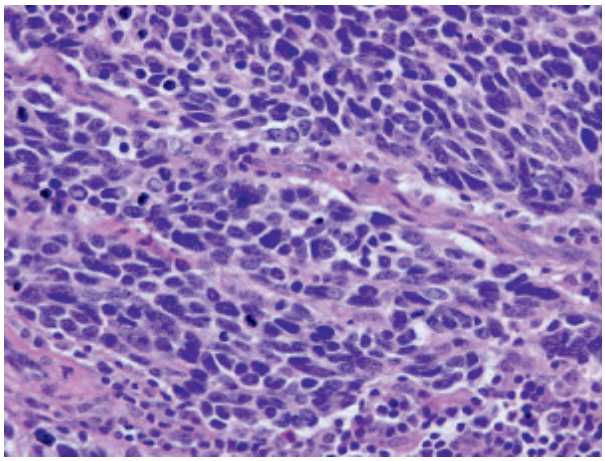


Figure 3: small cell carcinoma [29].

Non-small cell carcinoma

NSCLC are usually adenocarcinomas, squamous cell carcinomas (SQM) or large cell carcinomas (LCC) [29, 30, 32]. The distinguishing between the subtypes of NSCLC is necessary for the guidance of treatment and prediction of the clinical course.

Adenocarcinoma

Adenocarcinomas are the most common type of lung cancer, accounting for approximately half of all lung cancer cases [31, 38]. The incidence of adenocarcinoma is increasing, which is thought to be related to the introduction of low-tar filter cigarettes in the 1960s [38, 39]. The use of filters may have led to different inhalation behaviours, e.g., taking larger puffs and retaining smoke longer to compensate for the lower nicotine dose in the filter cigarettes. This might have led to increased carcinogenic damage in the peripheral lung zones, where the majority of the lung adenocarcinomas arise. Another explanation could be higher nitrate contents of the low-tar filter cigarettes. Also, lung cancer is increasing among females in whom adenocarcinoma seems to be more common [26].

In the case of adenocarcinoma, the tumour tissue is commonly tested for the presence of a driver mutation (e.g. mutated epidermal growth factor, ALK translocation) and increasingly for other mutations (**Figure 4**) [40]. This is necessary as it is possible to treat types of lung cancer in an advanced stage based on the genotype (so-called, personalised, genotype-directed therapy) [41, 42].

Bronchioalveolar carcinoma (BAC) develops in cells near the alveoli in the outer region of the lungs. This type of lung cancer is characterised by growth near the alveolar septae without evidence of stromal, vascular, or pleural invasion. Since 2016, the term BAC has ceased to be

used. Instead, the term "lepidic" is used, which describes non-invasive growth along intact septae [32]. Previous lesions classified as BAC are now classified under adenocarcinomas as:

I. Atypical adenomatous hyperplasia (AAH): ≤5mm. Previously recognised as pre-invasive lesion for lung adenocarcinoma.
II. Adenocarcinoma in situ (AIS): a localised adenocarcinoma, smaller than 3 cm, in which growth is restricted to tumour cells growing along alveolar structures (lepidic growth patter) and lacks any component of invasion. Most AIS are non-mucinous and just a small subset of such tumours are mucinous.
III. Minimally invasive adenocarcinoma: a new category which describes a small, solitary adenocarcinoma (≤3cm) with predominantly lepidic growth patterns and ≤5mm invasion.

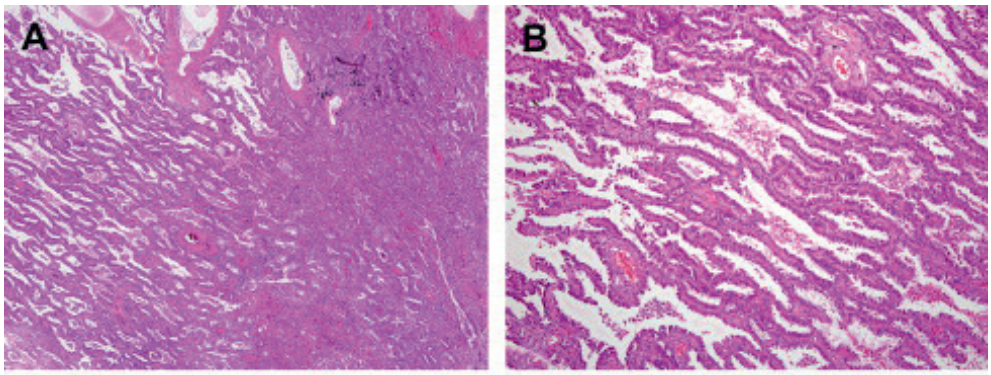


Figure 4: adenocarcinoma (A: Lepidic predominant pattern with mostly lepidic growth and B: invasive acinar adenocarcinoma) [29].

Squamous cell carcinoma

SQM used to be the most frequent type of lung cancer until the midst 1980s [29, 38]. Currently, it accounts for about 30% of all lung cancers [31, 38]. SQM arises in 60 to 80% in the proximal area of the tracheobronchial tree, although it is increasingly occurring as a peripheral lesion [43].

Together with SCLC, it shows a stronger correlation with cigarette smoking than LCC and adenocarcinomas [44, 45].

SQM most often arise in segmental bronchi, and involvement of lobar and main stem bronchus occurs by extension (**Figure 5**) [29]. In well-differentiated squamous cell carcinoma, squamous differentiation is suggested by intercellular bridging, squamous pearl formation and individual cell keratinisation which can readily be observed. In poorly differentiated squamous cell carcinoma, these features are difficult to find.

The diagnosis of SQM is mainly based on the presence of keratin production by the tumour cells and/or intercellular desmosomes, or by immunohistochemistry (expression of p40, p63, CK5, or CK5/6, desmoglein) [40, 46].

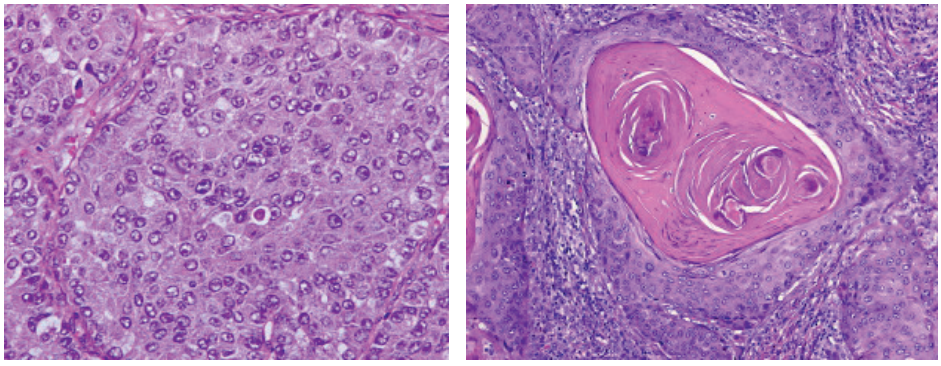


Figure 5: squamous cell carcinoma [29] and large cell carcinoma [29].

Large cell carcinoma

LCC comprises 3% of all lung cancers [1, 31]. The incidence of LCC is decreasing, due to reclassification of these tumours to mainly adenocarcinoma and squamous cell carcinoma [30]. These tumours are mostly found in the lung periphery, although they may have a central location. They frequently appear as large necrotic tumours.

LCC are malignant epithelial neoplasms lacking both glandular and squamous differentiation by light microscopy and immunohistochemistry, and lacking cytologic features of small cell carcinoma (**Figure 5**) [40, 47]. In other words, they are diagnosed by exclusion.

Other types of NSCLCs

Adenosquamous carcinomas are tumours which consist of at least 10% squamous cell carcinoma cells and at least 10% adenocarcinoma cells [48].

Carcinoids are neuroendocrine lung cancers with neuroendocrine differentiation lower than LCNEC and SCLC, and can further be divided into typical and atypical types [32]. This type of lung cancer can usually be diagnosed on the basis of light microscopy alone [29].

Non-small cell carcinoma, not otherwise specified (NSCLC NOS) is for cases where there is no evidence of squamous or adenocarcinomatous differentiation on immunohistochemistry and the endoscopic biopsies or cytology specimens are too small (sampling problems) [29].

Clinical manifestation

Unfortunately, the majority of patients is in advanced stage of disease at the time of lung cancer diagnosis. Symptoms of lung cancer do result from local effects of the tumour, from regional or distant spread, and/or from distant effects not related to metastases (e.g. paraneoplastic syndromes) [49-51].

Local effects of lung cancer:

I. Coughing: reported in more than 50% of the lung cancer patients at presentation and most frequently in those with a squamous cell and small cell carcinoma because of the central localization of the tumours [52-54].
II. Haemoptysis: reported in up to 30% of the patients with lung cancer [52-54].
III. Chest pain: reported in up to 40% of the patients with lung cancer [52-54].
IV. Dyspnoea: reported in approximately 50% of patients with lung cancer at presentation [52-54].
V. Other more general reported symptoms are: chest pain (20-49%), weight loss (27-68%), weakness (0-10%) and obstruction of the superior vena cava (0-4%) [52-55].

The most frequent sites of distant metastasis at the time of diagnosis or during the course of the disease are the liver, adrenal glands, bones and/or brain [56].

Lung cancer can also lead to symptoms which are mediated by hormones, cytokines or by an immune response against the tumour [36, 37]. Some commonly observed effects are: hypercalcemia (leading to anorexia, nausea, vomiting), syndrome of inappropriate antidiuretic hormone secretion (SIADH) in which the degree of hyponatremia leads to various symptoms, neurological paraneoplastic syndromes (e.g. difficulty to rise from a chair, dry mouth or stiff muscles), haematological symptoms such as thrombocytosis and Cushing's syndrome, in which an ectopic production of adrenal corticotropin (ACTH) is observed.

Treatment

Lung cancer is staged by using the International Association for the Study of Lung Cancer (IASCL) TNM staging system [57]. Lung cancers have been staged using the 8th edition of the TNM staging system, since 2017. However, lung cancers are described by the 7th edition in this thesis. Therefore the 7th TNM staging system is summarised and presented in **Table 1** and **Table 2**. The choice how to treat the patient with lung cancer depends on the TNM staging system [58].

In **Table 3** an overview is presented of the stage distribution at time of lung cancer diagnosis based upon the Surveillance, Epidemiology, and End Results (SEER) data of the United States between 2005 and 2014 [3].

Table 1: 7th edition TNM staging for lung cancer. Adapted from: Goldstraw *et al* (2007) [59].

Primary tumour (T)	
T1	Tumour ≤3 cm diameter in greatest dimension, surrounded by lung or visceral pleura, without invasion more proximal than lobar bronchus.
1a	Tumour ≤2 cm in diameter in greatest dimension.
1b	Tumour >2 cm but ≤3 cm in diameter in greatest dimension.
T2	Tumour >3 cm but ≤7 cm, or tumour with any of the following features: - Involves main bronchus, ≥2 cm distal to carina; - Invades visceral pleura; - Associated with atelectasis or obstructive pneumonitis that extends to the hilar region but does not involve the entire lung.
2a	Tumour >3 cm but ≤5 cm in greatest dimension.
2b	Tumour >5 cm but ≤7 cm in greatest dimension.
T3	Tumour >7 cm or any of the following: - Directly invades any of the following: chest wall, diaphragm, phrenic nerve, mediastinal pleura, parietal pericardium, main bronchus <2 cm from carina (without involvement of carina); - Atelectasis or obstructive pneumonitis of the entire lung; - Separate tumour nodules in the same lobe.
T4	Tumour of any size that invades the mediastinum, heart, great vessels, trachea, recurrent laryngeal nerve, esophagus, vertebral body, carina, or with separate tumour nodules in a different ipsilateral lobe.
Regional lymph nodes (N)	
N0	No regional lymph node metastases.
N1	Metastasis in ipsilateral peribronchial and/or ipsilateral hilar lymph nodes and intrapulmonary nodes, including involvement by direct extension.
N2	Metastasis in ipsilateral mediastinal and/or subcarinal lymph node(s).
N3	Metastasis in contralateral mediastinal, contralateral hilar, ipsilateral or contralateral scalene, or supraclavicular lymph node(s).
Distant metastasis (M)	
M0	No distant metastasis.
M1	Distant metastasis.
1a	Separate tumour nodule(s) in a contralateral lobe; tumour with pleural nodules or malignant pleural or pericardial effusion.
1b	Distant metastasis (in extra thoracic organs).

Table 2: stage grouping according to the seventh edition TNM staging. Adapted from: Goldstraw *et al* (2007) [59].

Stage groups	T stage	N stage	M stage
Ia	Ia,Ib	0	0
Ib	IIa	0	0
IIa	Ia, Ib	I	0
	IIa	I	0
	IIb	0	0
IIb	IIb	I	0
	III	0	0
IIIa	I-II	II	0
	III	I, II	0
	IV	0, I	0
IIIb	IV	II	0
	I-IV	III	0
IV	I-IV	0-III	Ia, Ib

Table 3: stage distribution of lung cancers at time of diagnosis [3].

Stages	Spread	%
Ia- IIb	Localised	18.2
IIIa-IIIb	Regional	21.9
IV	Distant	53.2
Unstaged	Unknown	6.7

Treatment of NSCLC

The NSCLCs are a heterogeneous group of lung cancers. In the case of curable disease, surgical resection offers the best opportunity for long-term survival of NSCLC patients. Molecular characterisation of tumour tissue serves as a guide to treatment in patients with metastatic disease and in those who relapse after primary treatment [60, 61].

In general, patients with stage I or stage II NSCLC are treated with surgical resection of the tumour combined with a surgical resection of the mediastinal lymph nodes [62, 63]. Surgical resection of the tumour is performed through a thoracotomy or increasingly by using video-assisted thoracoscopic surgery (VATS), a minimally invasive surgical method [64]. After an irradical resection of the tumour, or in the case of unexpected pathological N2 disease, adjuvant

radiotherapy is recommended [11]. For stage II resected tumours adjuvant chemotherapy is recommended [65]. For patients with compromised lung function or other co-morbidities (such as heart failure) who cannot undergo curative surgery for lung cancer, stereotactic radiation may be applied [66, 67].

Patients with a stage III NSCLC and a good clinical performance are treated with concurrent chemoradiation [62, 68]. Curative surgery plays a role in those patients with a down staging of the tumour after concurrent chemoradiation.

Patients with stage IV NSCLC, are treated with systemic therapy (e.g. chemotherapy) or symptom-based palliative approach (e.g. radiation for metastases) [62]. Therapy is guided by the mutation status of the tumour and by the clinical performance of the patient. Patients with metastasis may also benefit from resection of the metastasis (in case of isolated metastasis) as well as radiation (e.g. pain control). Local palliative measures (e.g. stenting of the vena cava, stenting of the oesophagus, coagulation of the bronchial vessels) may also be beneficial to control the pulmonary disease and increase the quality of life.

Table 4: overview of the NSCLC stage I or stage II treatment in the Netherlands [62].

Tumour status	**Type of treatment**
Tumour located in one lobe	Lobectomy with mediastinal lymph node resection
Tumour located in one lobe, but the patient does have a limited lung function	- Segmentation or wedge resection of the lobe with tumour - Or stereotactic radiotherapy
Tumour grows from one lobe into the next lobe	- Lobectomy of the main lobe with wedge resection of the next lobe - In case of central located tumour, bilobectomy or pneumonectomy
Tumour grows near by the central blood vessels	Sleeve lobectomy
Tumour grows in to the thorax wall	'En bloc' resection of the thorax wall next to surgical resection of the tumour
More than one tumour in one lobe	Lobectomy
More than one tumour in different lobes	Primary resection could be an option

Treatment of SCLC

At the time of diagnosis, SCLC is usually disseminated (extensive). As SCLC is very responsive to chemotherapy, systemic chemotherapy is the main treatment for SCLC [69].

In general, patients with a limited (not extensive) SCLC are treated with a combination

of chemotherapy and radiotherapy [62, 70]. Surgery is only used in patients with a solitary pulmonary nodule without metastasis or lymph node involvement [71]. Patients without disease progression and with a good clinical performance are recommended to receive a prophylactic cranial radiation (PCI) within 60 days after chemotherapy [72].

For patients with extensive SCLC, chemotherapy is used as the only initial therapy [62]. PCI is considered in patients with a complete or partial response, or stable disease after the chemotherapy. Radiation therapy can also be used for symptom-based control.

Survival

In general, the 5-year survival rate for lung cancer is about 11-18% [3, 5, 31]. Survival decreases progressively with later stages of the disease. In **Table 5** an overview is presented of the 5-year survival rate based upon the Surveillance, Epidemiology, and End Results (SEER) data of lung cancer survival in the United States between 2007 and 2013 [3].

In an analysis of surgical NSCLC cases from the IASLC lung cancer staging project, next to stage, age and gender were prognostic factors for survival [73]. Patients less than 70 years of age had significantly a better overall survival compared to patients older than 70 years of age. BACs carried a better prognosis than all other subtypes of lung cancer. Squamous cell carcinomas were slightly favoured over adenocarcinomas and LCC, but only after adjusting for gender (male only) and stage. Adjusting for smoking status did not modify the effects of histology.

Furthermore, histologic grade has significant a prognostic value for the survival of NSCLC lung cancer patients, in which undifferentiated carcinoma have an elevated risk of death compared to well-differentiated and moderate-differentiated carcinoma [74]. The majority of lung cancer patients who died perioperative are current smokers [75, 76]. Non-smokers, former smokers and recent quitters do have a significantly better prognosis than current smokers with lung cancer [75, 77].

Table 5 : five-year survival rate (%) of NSCLC and SCLC by stage [3].

Stages	**Spread**	**NSCLC**	**SCLC**
Ia- IIb	Localised	59.5	28.9
IIIa-IIIb	Regional	32.3	15.3
IV	Distant	5.2	2.9
Unstaged	Unknown	13.4	7.6
All stages	-	22.1	6.5

NSCLC: non-small cell carcinoma; SCLC: small cell carcinoma.

Lung cancer screening

The majority of lung cancers are clinically diagnosed at an advanced disease stage. As a result, treatment options are limited which leads to a low 5-year lung cancer survival rate [3, 5, 31].

Preventing lung cancer by controlling the risk factors for lung cancer is called 'primary prevention for lung cancer'. As approximately 85% of lung cancer cases are related to tobacco smoking, the abstinence of smoking and prevention of initiating the use of tobacco leads to a decrease in lung cancer incidence [1]. Currently, in the United States 15.1% of the population is a prevalent smoker and in the Netherlands 26.3% of the population older than 18 years currently smokes [11, 12, 78]. In low-income countries the prevalence of smoking is increasing [2, 31, 49, 79]. Smoking cessation is an important prevention method for decreasing the incidence of lung cancer in the long term [80, 81]. However, the efficacy of the current smoking cessation programmes aimed at the general population is insufficient [82, 83]. Moreover, increasingly lung cancers are being diagnosed in former smokers [84, 85]. This underscores the need for early detection and treatment (secondary prevention) of lung cancer.

Secondary prevention (screening) is detection of pre-clinical lung cancer lesions in asymptomatic persons, aiming to increase the opportunities for treatment and prevent progression of the cancer. The target population are people who are at high risk for developing lung cancer, but who are not already diagnosed with the disease.

Before the 1990s, chest X-ray and sputum cytology were studied in clinical trials as a potential screening test for lung cancer and showed no significant lung cancer mortality reduction [86-89]. By the introduction of low-dose Computed Tomography (CT) scanning in the 1990s a new period started of investigating CT scanning as a screening test for lung cancer. Different single-arm studies of lung cancer screening with CT scanning showed that more lung cancers can be detected in an early stage using this method [90-94]. However, to overcome various biases, such as, lead-time, length-time and overdiagnosis (discussed in more depth below), a randomised-controlled setting with a control group without intervention was necessary.

The largest randomised-controlled lung cancer CT screening trial is the National Lung Screening Trial (NLST), which took place in the United States from 2002 until 2011 [95, 96]. The primary aim was to investigate whether low-dose Computed Tomography (LDCT) leads to a lung cancer mortality reduction of at least 20% in subjects at a high risk for developing lung cancer. In total, 53,434 current or former smokers (who had smoked at least 30 pack years and did not quit smoking more than 15 years ago) aged between 55 and 74 years without symptoms or signs of lung cancer were enrolled. Participants received three annual screenings either with LDCT (screen group) or with standard chest radiography (control group). In 2011, the study demonstrated that LDCT screening led to a lung cancer mortality reduction in high-risk subjects by 18 to 20% compared to chest radiography screening. Fewer deaths per 1,000 screened participants in the LDCT group were observed compared to the chest radiography group over an average of 6.5 years of follow-up (17.6 per 1,000 vs. 20.7 per 1000,

respectively) [96, 97]. Furthermore, the all-cause mortality was 6.7% lower in the LDCT screening group compared to the chest radiography group. The number of scanned individuals required to prevent 1 lung cancer death was 320 and the number to prevent 1 death overall was 219 over 6.5 years.

Based on the positive outcomes of the NLST, the United States Preventive Services Task Force (USPSTF) requested an independent review and a comparative modelling study to investigate the effectiveness of lung cancer screening using CT scans [98, 99]. In 2013, this led to a grade B recommendation of annual CT lung cancer screening in subjects aged between 55 and 80 and who have a 30 pack year smoking history and who are current smokers or quit in the past 15 years [96, 98]. A grade B recommendation is suggested for practice, which means that there is a high certainty that the net benefit is moderate or there is a moderate certainty that the net benefit is moderate to substantial.

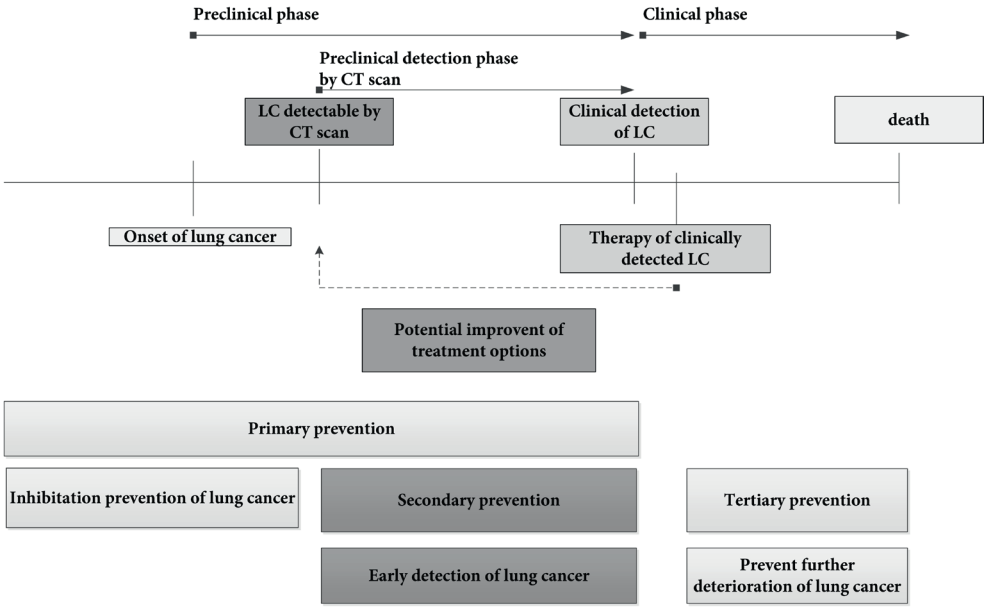


Figure 6: prevention stages of lung cancer.
CT: Computed Tomography. LC: lung cancer.

In Europe, seven randomised-controlled trials of LDCT have been conducted or are still in progress at this moment (**Table 6**) [100-106]. The main difference with the NLST is that the European trials all used a control group in which no screening was offered. So far, four European trials have shown no significant lung cancer mortality reduction (**Table 7**), but it should be noted that these trials were strongly underpowered (e.g. small sample size) to be able to show a possible benefit. The mortality analyses of the NELSON, LUSI and UKLS studies are expected in the coming years.

Table 6: an overview of the seven European CT lung cancer screening trials and their study protocol.

	n	**Recruitment method**	**Age group**	**Smoking history**	**Smoking cessation**	**Study groups**	**Screening interval**
NELSON [100, 107]	15,792	Population-based	50-75	>15 cigarettes p/d for >25 years or >10 cigarettes p/d for >30 years	≤10 years	CT vs usual care	Four screenings with screening intervals of 1, 2, and 2,5 years
DLCST [101]	4,104	Volunteers	50-70	≥20 pack-years	≤10 years	CT vs usual care	Five annual screenings
MILD [102]	4,099	Volunteers	>49	≥20 pack-years	≤10 years	CT vs usual care	Five annual screenings vs three biennial screenings
UKLS [103]	4,061	Population-based	50-75	5-year lung cancer risk, and predicted risk of ≥5% of lung cancer diagnosis within 5 years	≤10 years	CT vs usual care	One screening
LUSI [104]	4,052	Population-based	50-69	≥15 cigarettes p/d for ≥25 years or ≥10 cigarettes p/d for ≥30 years	≤10 years	CT vs usual care	Four annual screenings

Table 6: continued. an overview of the seven European CT lung cancer screening trials and their study protocol.

ITALUNG [105]	3,206	Through general practitioners	55-69	≥20 pack-years	≤10 years	CT vs usual care	Four annual screenings
DANTE [106]	2,472	Volunteers	60-74	≥20 pack-years	≤10 years	CT vs baseline chest radio-graphy	Four annual screenings

NELSON: The Dutch-Belgian Lung cancer Screening Trial; DLCST: Danish Lung Cancer Screening Trial; MILD: Multicentric Italian Lung Detection trial; UKLS: UK Lung cancer Screening pilot trial; LUSI: Lung tumour screening and intervention trial; ITALUNG; The Italian Lung Study; DANTE: Detection And screening of early lung cancer with Novel imaging Technology.

Table 7: an overview of the seven European CT lung cancer screening trials and their end-point analysis.

	NELSON	**DLCST [108, 109]**	**MILD [99]**	**UKLS**	**LUSI**	**ITALUNG [110]**	**DANTE [111]**
LC mortality, RR (95%CI)	-	1.03 (0.66-1.60)	1.64[1] (0.73-4.01)	-	-	0.70 (0.47-1.03)	0.99 (0.69-1.43)
All-cause mortality, RR (95%CI)	-	1.02 (0.82-1.27)	1.40[2] (0.82-2.38)	-	-	0.83 (0.67-1.03)	0.95 (0.77-1.17)

[1]There was no significant difference between the annual and biennial screening arm (p=0.21). This presents the HR when comparing the two LDCT arms together with the control group; [2]There was no significant difference between the annual and biennial screening arm (p=0.13). This presents the HR when comparing the two LDCT arms together with the control group

Criteria for lung cancer screening programme

To justify a screening programme it should meet the following requirements:
1) there should be substantial positive health outcomes (e.g. mortality reduction, life years gained, significant increase in management or treatment options); 2) the adverse side-effects (e.g. extent of early detection, overdiagnosis and side-effects) should be limited; and 3) the ratio between costs and benefits should be reasonable.

The most important benefit of lung cancer screening is the reduction in lung cancer mortality (**Table 8**). This has been demonstrated by the NLST and has been an important factor in the implementation of LDCT screening for lung cancer in the United States [98]. The lung cancer mortality rate is widely accepted as a measurement to determine the effectiveness of a lung cancer screening programme. Case-specific survival of lung cancer (or survival rate) is not recommended since it refers to the number of people with lung cancer remaining alive at a certain point in time after diagnosis. Moreover, survival rate does not adjust for the effects of lead-time, length-time and overdiagnosis biases (**Figure 7 and Figure 8**):

- Lead-time bias: length of time between early diagnosis of cancer by screening and the time in which the diagnosis would have been made without screening. In this case, the survival time is increased without affecting the actual course of the disease. In other words, there is no benefit of screening.
- Overdiagnosis bias: detection of tumours by screening, which may have remained subclinical before death from another cause.
- Length-time bias: screening is most likely to detect relatively slow-growing tumours, because they have a longer interval of being visible to be detectable by CT scan and have a longer asymptomatic phase. It gives the appearance that screening prolongs life. Therefore, in mortality analyses mortality of both the screen group and control group should be investigated.

Other benefits of lung cancer screening are the reduction in all-cause mortality and increase of early staged lung cancers with more favourable treatment options [96, 112]. Furthermore, participation in LDCT screening may be a teachable moment to quit smoking [113, 114].

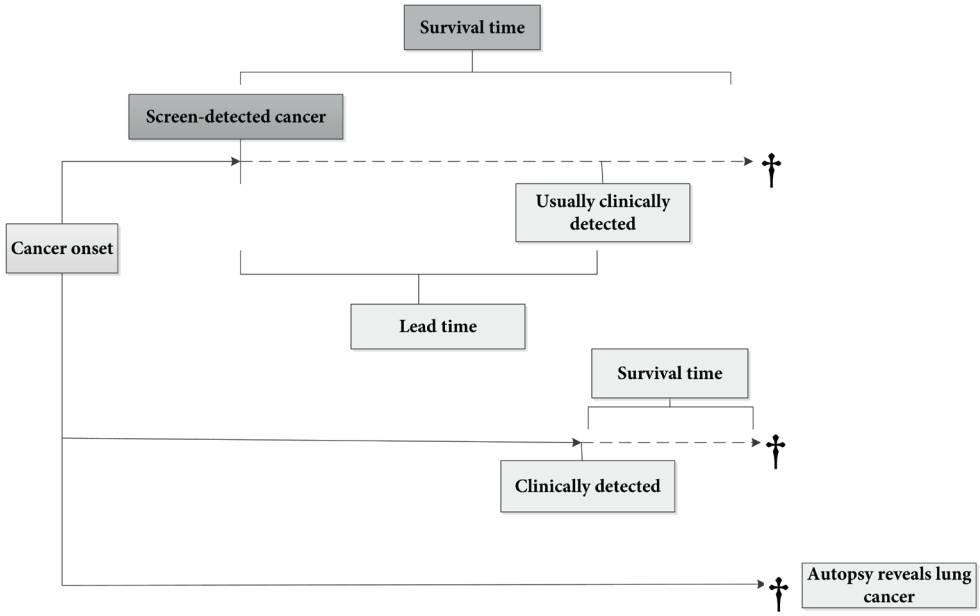


Figure 7: lead-time bias and overdiagnosis bias.

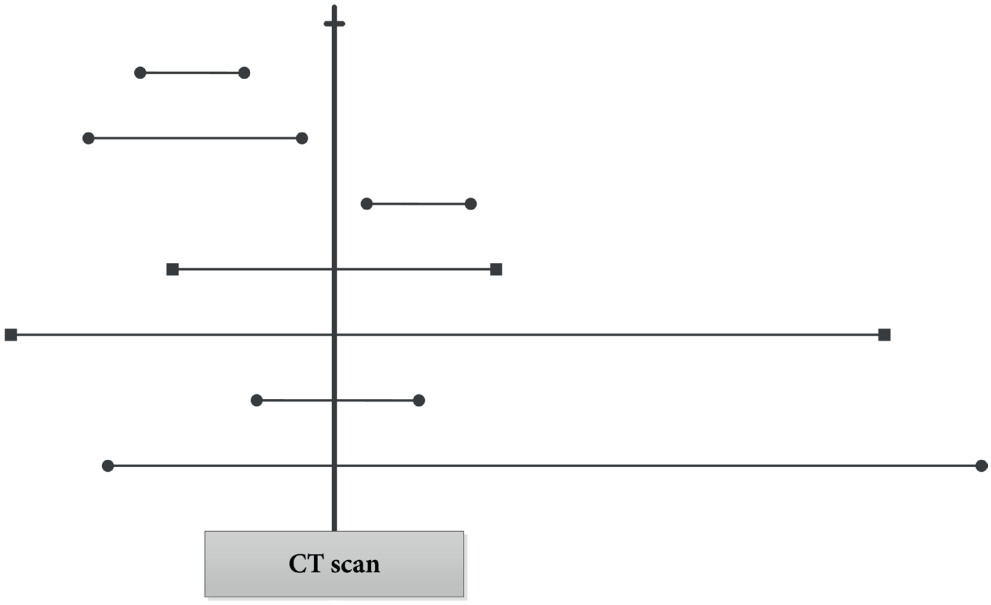


Figure 8: length time bias. CT: computed tomography.
Bullet: fast-growing tumours. Squares: slow-growing tumours.

However, screening might unintentionally expose the screened population to a variety of harms. One of the primary potential harm is a false positive test result: a benign nodule identified by CT-scanning which may leads to unnecessary invasive investigation. The false-positive rate across the three screening rounds in the NLST was quite high (23.3%) [96]. Another example is the delay of lung cancer diagnosis by a false negative test result [115]. Moreover, screening will also lead to detection of lung cancers which may have not affected the patient's lifetime if left untreated (overdiagnosis). True extent of overdiagnosis in lung cancer is difficult to determine because most of the current information what we know about lung cancer is derived from symptomatic patients. One study using excess-incidence reported that more than 18% of all lung cancers detected by LCDT in the NLST are indolent [116]. While, another study using Microsimulation SCreening Analysis (MISCAN) Lung model with the NLST data estimated that at year eight overdiagnosis rate to be 12.5% compared to 18.5%[117]. Hence, it is important to maximise the benefits while simultaneously minimizing the harms.

Table 8: benefits and harms of lung cancer screening.

Benefits	**Harms**
Overall mortality and lung cancer mortality reduction	Complications of screen result and diagnostic work-up (false positives, invasive diagnostic work-up associated with morbidity and mortality)
Reduction in lung cancer incidence and advanced stage lung cancers	Delay of lung cancer diagnosis by a false negative test result
DALY/QALY/life-years gained	Overdiagnosis
Increase of curative treatment options	Radiation exposure (induces the risk for developing cancer)
A teachable moment for smoking cessation	Psychological consequences (patient distress, anxiety)
Improvement in diagnostic procedures and cancer treatment	Possible negative effect on smoking cessation and false reassurance

DALY=disability-adjusted life-year; QALY=quality-adjusted life-year

The NELSON trial

Study design

The largest European lung cancer screening trial, the NELSON trial was initiated in the Netherlands and in Belgium in 2003 [107].

The primary aims of this trial are:

I. To establish whether LDCT screening in high-risk subjects would lead to a reduction of ≥25% in lung cancer mortality;
II. To estimate the impact of lung cancer screening on health-related quality of life and smoking cessation;
III. To estimate cost-effectiveness of LDCT screening for lung cancer.

Study participants were randomised (1:1) into no screening (control group) or screening using LDCT. Screening took place at baseline (round 1), after one year (round 2), after three years (round 3) and after five and half years (round 4). Participants of both groups are followed-up and the difference in lung cancer mortality between the study groups will be determined ten years after randomisation (Figure 9).

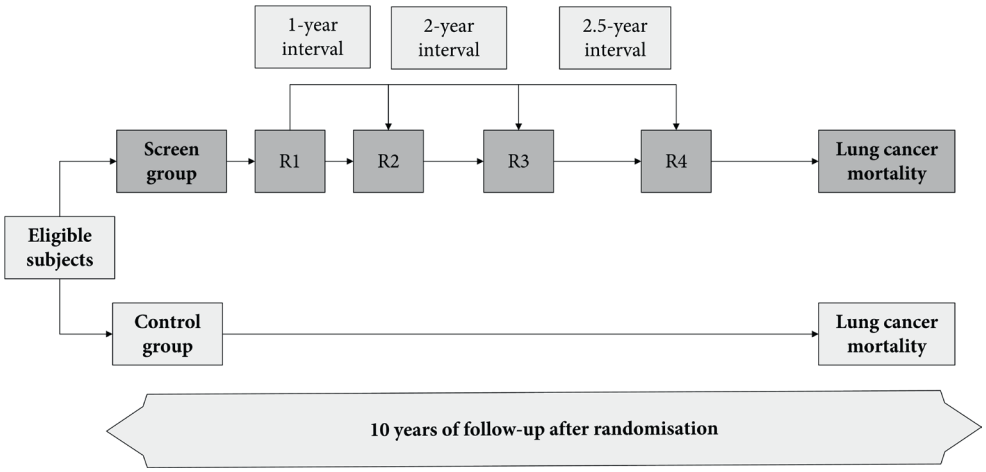


Figure 9: study design of the NELSON trial.

Population-based recruitment method

To minimise the "self-selection bias" that occurs when healthy volunteers are more eager to participate in a screening programme, a population-based recruitment strategy was used.

Between the second half of 2003 and the second half of 2005, addresses of approximately 600,000 subjects aged 50-74 years were obtained from seven districts in the Netherlands and 14 municipalities around Leuven (Belgium) [100]. These subjects received a questionnaire about their general health status, medical check-ups and history, physical activity, body weight and length, smoking history, alcohol consumption, their own medical history and family history of cancer, level of education and their opinions on screening programmes [100, 118]. The questionnaire did not contain any information about the upcoming lung cancer screening trial.

The information obtained from this first questionnaire was used to decide who met the inclusion- and/or exclusion criteria of the trial (**Table 9**) [100]. Respondents who were eligible received an information leaflet about the NELSON trial, an invitation to participate, an informed consent form for participating in the NELSON trial and a second questionnaire (about smoking habits and exposure to asbestos in more detail).

Finally, those subjects who provided the informed consent and completed the second questionnaire (initially n=15,822, later adjusted to 15,792) were randomised (1:1) to either the screening group or the control group.

The inclusion criteria were:

I. Aged between 50 and 74 at randomization;
II. Smoking history >15 cigarettes per day during >25 years or >10 cigarettes per day during >30 years;
III. Smoking cessation ≤ 10 years ago.

The exclusion criteria were:

I. A moderate or poor self-reported health who were unable to climb two flights of stairs;
II. A body weight ≥140 kilogram;
III. A lung cancer diagnosis <5 years ago or ≥5 years ago but still under treatment;
IV. A current or past renal cancer, breast cancer or melanoma;
V. A CT chest examination <1 year ago;
VI. A uncompleted informed consent.

Screening procedure and protocol

Screen group participants were invited by mail for an LDCT scan at one of the four screening sites: University Medical Centre Groningen, University Medical Centre Utrecht and Kennemer Gasthuis Haarlem in the Netherlands or at University Hospital Gasthuisberg Leuven in Belgium [107, 119]. For the LDCT examination, 16-detector, or in later rounds 64-detector CT scanners in low-dose settings were used, without the administration of intravenous contrast media [120]. Data acquisition and scanning conditions were kept standard across the four screening centres for the duration of the trial.

Images were analysed using semi-automatic software (LungCARE, version Somaris/5 VA70C-W, Siemens Medical Solutions), in which the semi-automatic segmentation of nodules and determination of the nodule volume took place. In case the software could not segment the nodule accurately, the radiologist measured the diameter manually [107].
A nodule was defined as a small, spherical, non-linear circumscribed focus of abnormal tissue. The nodule characteristics (e.g. diameter, volume, density, location, lung segment and surface characteristics as smooth, speculated or other) were recorded and uploaded immediately in the NELSON Nodule Management System (NMS) [107].

In the first two rounds, two radiologists independently reviewed the images in NMS. In case of a discrepancy, a third expert reader made the final decision [107, 120]. However, for the last two rounds a single reading was performed, as Wang et al. showed that there was no benefit for double reading consensus with the use of semi-automated software [121].

The screening test had three possible results, depending on the presence of one or more nodule(s), nodule volume and volume doubling time (VDT) [119]:

I. Negative: no nodule detected, screen result not suspicious for lung cancer. No further diagnostic tests warranted;
II. Indeterminate: a small abnormality identified for which at this moment no further investigation is needed, however, in order to see whether there has been a change in this abnormality over time a follow-up scan will be made;
III. Positive: abnormalities suspicious for lung cancer. Participant referred to a pulmonologist for a diagnostic work-up.

In the case of newly detected solid nodules and the solid component of the part-solid nodules, the volume determined the screening outcome. In the case of previously detected nodules, the volume growth (change in volume of the nodule) and the VDT of the nodule determined the outcome.

I. Negative screening outcome: a newly detected nodule volume of $<50mm^3$ or, in the case of a previously detected nodule, volume growth of ≤25% or a volume growth of >25% and a VDT of >600 days. These participants were invited for the next screening round. For those in the fourth round, this was the end of the screening programme;
II. Indeterminate screening outcome: a newly detected nodule volume of $50mm^3$-$500mm^3$ or, in the case of a previously detected nodule, a volume growth of >25% and a VDT of 400-600 days. These participants received a follow-up LDCT examination according to protocol (after 6-12 weeks, or after 3-4 months) to classify the final result as positive or negative;
III. Positive screening outcome: a newly detected nodule volume of $>500mm^3$ or, in the case of a previously detected nodule a volume growth of >25% and a VDT of <400 days. These participants were immediately referred to a pulmonologist for a diagnostic work-up. If the work-up did not lead to the diagnosis of lung cancer, they were invited

for the next screening round. For those in case of the fourth round, this was the end of the screening programme.

All lung cancers were staged using the 7th TNM staging system [57, 59].

Data collection

Prospectively, all relevant medical data about the participants in the two groups is being collected. Relevant medical data contains information about the diagnosis, treatment and follow-up of the lung cancer, and about the cause of death. Furthermore, it contains information about the diagnostic work-up performed in the participants with positives screen test results.

Information about the participants from whom all relevant medical data needs to be collected, is obtained through data linkages with the Dutch Cancer Registry (NKR) and the Belgian Cancer Registry.

After collecting medical data it was first verified whether the participant was indeed diagnosed with lung cancer during the course of the study or at the time of autopsy.

End point verification

Linkages with Statistics Netherlands provided the cause of death of deceased participants. However, a clinical expert death review committee has been formed in the NELSON trial in order to accurately verify the cause of death of the study participants. The committee consists of two main reviewers: a pulmonologist-oncologist and a pathologist specialised in lung oncology. Only in case when no consensus is reached between the two reviewers a third reviewer (an epidemiologist specialised in screening) is consulted. This committee reviews the blinded medical files of the deceased participants diagnosed with lung cancer, independently and in a uniform way by using a death review process protocol [122].

Research questions

In this thesis three research topics are addressed: 1) the optimization of the NELSON screening rounds; 2) the implications and generalizability of the NELSON trial results; and 3) the cause of death verification process of the deceased NELSON study participants

Part I: the optimization of the NELSON screening rounds

I. What is the added value of a fourth screening round with an interval of 2.5 years after the previous three screening rounds?

II. Which NELSON subgroups with different risks for detecting lung cancer can be identified based on their previous screening history?

Part II: interim stage shift results in the NELSON trial

III. What is the level of cancer and treatment shift between the two study groups in a selection of the study cohort?

Part III: the cause of death of the NELSON study participants

IV. What are the differences in characteristics and mortality profile of NELSON participants and eligible non-responders?

V. What is the outcome of the Cause of Death verification process in the NELSON trial, and how does it relate to the official death certificates?

Outline of this thesis

The research questions are divided into three parts. Part 1 consists of two chapters that address the "Optimization of the NELSON screening rounds" by determining the added value of a fourth screening round (**Chapter II**), and in which NELSON subgroups with various risks for detecting lung cancer can be identified based on their previous screening results (**Chapter III**). Part 2 consists of one chapter, which presents the interim stage shift results of the NELSON trial (**Chapter IV**). Part 3 addresses the difference in characteristics and mortality profile of the NELSON participants and eligible non-responders (**Chapter V**) and makes a comparison between the cause of death verification process in the NELSON trial and the official death certificates (**Chapter VI**). This is essential for the primary end results of this trial (lung cancer specific mortality).

A general discussion is presented in **Chapter VII**, in which the published articles referred to in this thesis are reviewed in order to interpret important results, answer the research questions and to formulate general conclusions and recommendations.

References

1. Torre LA, Siegel RL, Jemal A. Lung Cancer Statistics. Adv Exp Med Biol. 2016;893:1-19.
2. Islami F, Torre LA, Jemal A. Global trends of lung cancer mortality and smoking prevalence. Transl Lung Cancer Res. 2015;4(4):327-38.
3. Howlader N, Noone AM, Krapcho M, Miller D, Bishop K, Kosary CL, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA (eds). SEER Cancer Statistics Review, 1975-2014, National Cancer Institute. Bethesda, MD, https://seer.cancer.gov/csr/1975_2014/, based on November 2016 SEER data submission, posted to the SEER web site, April 2017. Access date 19-07-2017.
4. Dataset regarding lung cancer deaths provided by Statistics Netherlands (CBS), publically accesable from statline.cbs.nl, access date 04-04-2017.
5. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2017. CA Cancer J Clin. 2017;67(1):7-30.
6. Wynder EL, Graham EA. Landmark article May 27, 1950: Tobacco Smoking as a possible etiologic factor in bronchiogenic carcinoma. A study of six hundred and eighty-four proved cases. By Ernest L. Wynder and Evarts A. Graham. Jama. 1985;253(20):2986-94.
7. Wynder EL, Graham EA. Tobacco smoking as a possible etiologic factor in bronchiogenic carcinoma; a study of 684 proved cases. J Am Med Assoc. 1950;143(4):329-36.
8. Bach PB. Smoking as a factor in causing lung cancer. Jama. 2009;301(5):539-41.
9. Thun M, Peto R, Boreham J, Lopez AD. Stages of the cigarette epidemic on entering its second century. Tob Control. 2012;21(2):96-101.
10. Netherlands Comprehensive Cancer Organisation, IKNL. Publicaly accessible data available on www cijfersoverkanker nl, 2016. Access date 19-07-2017 by www.cijfersoverkanker.nl
11. Billiet C, Decaluwe H, Peeters S, Vansteenkiste J, Dooms C, Haustermans K, et al. Modern post-operative radiotherapy for stage III non-small cell lung cancer may improve local control and survival: a meta-analysis. Radiother Oncol. 2014;110(1):3-8.
12. van Laar MW, Nationale Drug Monitor. Published in 2016 by Trimbos Institute. Accesable from www.trimbos.nl/kerncijfers/nationale-drug-monitor. Access date 23-08-2017.
13. Alberg AJ, Samet JM. Epidemiology of lung cancer. Chest. 2003;123(1 Suppl):21S-49S.
14. Pesch B, Kendzia B, Gustavsson P, Jockel KH, Johnen G, Pohlabeln H, et al. Cigarette smoking and lung cancer--relative risk estimates for the major histological types from a pooled analysis of case-control studies. Int J Cancer. 2012;131(5):1210-9.
15. Villeneuve PJ, Mao Y. Lifetime probability of developing lung cancer, by smoking status, Canada. Can J Public Health. 1994;85(6):385-8.
16. Jemal A, Ward E, Hao Y, Thun M. Trends in the leading causes of death in the United States, 1970-2002. Jama. 2005;294(10):1255-9.
17. Spitz MR, Hong WK, Amos CI, Wu X, Schabath MB, Dong Q, et al. A risk model for prediction of lung cancer. J Natl Cancer Inst. 2007;99(9):715-26.
18. Matakidou A, Eisen T, Houlston RS. Systematic review of the relationship between family history and lung cancer risk. Br J Cancer. 2005;93(7):825-33.
19. Brenner DR, McLaughlin JR, Hung RJ. Previous lung diseases and lung cancer risk: a systematic review and meta-analysis. PLoS One. 2011;6(3):e17479.

20. Yu YH, Liao CC, Hsu WH, Chen HJ, Liao WC, Muo CH, et al. Increased lung cancer risk among patients with pulmonary tuberculosis: a population cohort study. J Thorac Oncol. 2011;6(1):32-7.
21. Young RP, Hopkins RJ, Christmas T, Black PN, Metcalf P, Gamble GD. COPD prevalence is increased in lung cancer, independent of age, sex and smoking history. Eur Respir J. 2009;34(2):380-6.
22. Besaratinia A, Pfeifer GP. Second-hand smoke and human lung cancer. Lancet Oncol. 2008;9(7):657-66.
23. Brennan P, Buffler PA, Reynolds P, Wu AH, Wichmann HE, Agudo A, et al. Secondhand smoke exposure in adulthood and risk of lung cancer among never smokers: a pooled analysis of two large studies. Int J Cancer. 2004;109(1):125-31.
24. Villeneuve PJ, Parent ME, Harris SA, Johnson KC, Canadian Cancer Registries Epidemiology Research G. Occupational exposure to asbestos and lung cancer in men: evidence from a population-based case-control study in eight Canadian provinces. BMC Cancer. 2012;12:595.
25. Darby S, Hill D, Auvinen A, Barros-Dios JM, Baysson H, Bochicchio F, et al. Radon in homes and risk of lung cancer: collaborative analysis of individual data from 13 European case-control studies. Bmj. 2005;330(7485):223.
26. Kligerman S, White C. Epidemiology of lung cancer in women: risk factors, survival, and screening. AJR Am J Roentgenol. 2011;196(2):287-95.
27. Lorigan P, Radford J, Howell A, Thatcher N. Lung cancer after treatment for Hodgkin's lymphoma: a systematic review. Lancet Oncol. 2005;6(10):773-9.
28. Kirova YM, Gambotti L, De Rycke Y, Vilcoq JR, Asselain B, Fourquet A. Risk of second malignancies after adjuvant radiotherapy for breast cancer: a large-scale, single-institution review. Int J Radiat Oncol Biol Phys. 2007;68(2):359-63.
29. Travis WD. Pathology of lung cancer. Clin Chest Med. 2011;32(4):669-92.
30. Travis WD, Brambilla E, Noguchi M, Nicholson AG, Geisinger KR, Yatabe Y, et al. International association for the study of lung cancer/american thoracic society/european respiratory society international multidisciplinary classification of lung adenocarcinoma. J Thorac Oncol. 2011;6(2):244-85.
31. Torre LA, Siegel RL, Jemal A. Lung Cancer Statistics. Adv Exp Med Biol. 2016;893(0065-2598 (Print)):1-19.
32. Travis WD, Brambilla E, Nicholson AG, Yatabe Y, Austin JH, Beasley MB, et al. The 2015 World Health Organization Classification of Lung Tumors: Impact of Genetic, Clinical and Radiologic Advances Since the 2004 Classification. J Thorac Oncol. 2015;10(9):1243-60.
33. Karachaliou N, Pilotto S, Lazzari C, Bria E, de Marinis F, Rosell R. Cellular and molecular biology of small cell lung cancer: an overview. Transl Lung Cancer Res. 2016;5(1):2-15.
34. Kitamura H, Yazawa T, Sato H, Okudela K, Shimoyamada H. Small cell lung cancer: significance of RB alterations and TTF-1 expression in its carcinogenesis, phenotype, and biology. Endocr Pathol. 2009;20(2):101-7.
35. Kasprzak A, Przewozna M, Surdyk-Zasada J, Zabel M. The expression of selected neuroendocrine markers and of anti-neoplastic cytokines (IL-2 and IL-12) in lung cancers. Folia Morphol (Warsz). 2003;62(4):497-9.

36. Bentea G, Sculier C, Grigoriu B, Meert AP, Durieux V, Berghmans T, et al. Autoimmune paraneoplastic syndromes associated to lung cancer: A systematic review of the literature: Part 3: Neurological paraneoplastic syndromes, involving the central nervous system. Lung Cancer. 2017;106:83-92.
37. Richardson GE, Johnson BE. Paraneoplastic syndromes in lung cancer. Curr Opin Oncol. 1992;4(2):323-33.
38. Janssen-Heijnen ML, Coebergh JW. The changing epidemiology of lung cancer in Europe. Lung Cancer. 2003;41(3):245-58.
39. Janssen-Heijnen ML, Coebergh JW, Klinkhamer PJ, Schipper RM, Splinter TA, Mooi WJ. Is there a common etiology for the rising incidence of and decreasing survival with adenocarcinoma of the lung? Epidemiology. 2001;12(2):256-8.
40. Rekhtman N, Ang DC, Sima CS, Travis WD, Moreira AL. Immunohistochemical algorithm for differentiation of lung adenocarcinoma and squamous cell carcinoma based on large series of whole-tissue sections with validation in small specimens. Mod Pathol. 2011;24(10):1348-59.
41. Kumarakulasinghe NB, van Zanwijk N, Soo RA. Molecular targeted therapy in the treatment of advanced stage non-small cell lung cancer (NSCLC). Respirology. 2015;20(3):370-8.
42. Hensing T, Chawla A, Batra R, Salgia R. A personalized treatment for lung cancer: molecular pathways, targeted therapies, and genomic characterization. Adv Exp Med Biol. 2014;799:85-117.
43. Funai K, Yokose T, Ishii G, Araki K, Yoshida J, Nishimura M, et al. Clinicopathologic characteristics of peripheral squamous cell carcinoma of the lung. Am J Surg Pathol. 2003;27(7):978-84.
44. Khuder SA. Effect of cigarette smoking on major histological types of lung cancer: a meta-analysis. Lung Cancer. 2001;31(2-3):139-48.
45. Yun YH, Lim MK, Jung KW, Bae JM, Park SM, Shin SA, et al. Relative and absolute risks of cigarette smoking on major histologic types of lung cancer in Korean men. Cancer Epidemiol Biomarkers Prev. 2005;14(9):2125-30.
46. Kim MJ, Shin HC, Shin KC, Ro JY. Best immunohistochemical panel in distinguishing adenocarcinoma from squamous cell carcinoma of lung: tissue microarray assay in resected lung cancer specimens. Ann Diagn Pathol. 2013;17(1):85-90.
47. Rossi G, Mengoli MC, Cavazza A, Nicoli D, Barbareschi M, Cantaloni C, et al. Large cell carcinoma of the lung: clinically oriented classification integrating immunohistochemistry and molecular biology. Virchows Arch. 2014;464(1):61-8.
48. Filosso PL, Ruffini E, Asioli S, Giobbe R, Macri L, Bruna MC, et al. Adenosquamous lung carcinomas: a histologic subtype with poor prognosis. Lung Cancer. 2011;74(1):25-9.
49. Kocher F, Hilbe W, Seeber A, Pircher A, Schmid T, Greil R, et al. Longitudinal analysis of 2293 NSCLC patients: a comprehensive study from the TYROL registry. Lung Cancer. 2015;87(2):193-200.
50. Buccheri G, Ferrigno D. Lung cancer: clinical presentation and specialist referral time. Eur Respir J. 2004;24(6):898-904.
51. Hamilton W, Peters TJ, Round A, Sharp D. What are the clinical features of lung cancer before the diagnosis is made? A population based case-control study. Thorax. 2005;60(12):1059-65.

52. Hyde L, Hyde CI. Clinical manifestations of lung cancer. Chest. 1974;65(3):299-306.
53. Chute CG, Greenberg ER, Baron J, Korson R, Baker J, Yates J. Presenting conditions of 1539 population-based lung cancer patients by cell type and stage in New Hampshire and Vermont. Cancer. 1985;56(8):2107-11.
54. Spiro SG, Gould MK, Colice GL, American College of Chest P. Initial evaluation of the patient with lung cancer: symptoms, signs, laboratory tests, and paraneoplastic syndromes: ACCP evidenced-based clinical practice guidelines (2nd edition). Chest. 2007;132(3 Suppl):149S-60S.
55. Eren S, Karaman A, Okur A. The superior vena cava syndrome caused by malignant disease. Imaging with multi-detector row CT. Eur J Radiol. 2006;59(1):93-103.
56. Riihimaki M, Hemminki A, Fallah M, Thomsen H, Sundquist K, Sundquist J, et al. Metastatic sites and survival in lung cancer. Lung Cancer. 2014;86(1):78-84.
57. Marshall HM, Leong SC, Bowman RV, Yang IA, Fong KM. The science behind the 7th edition Tumour, Node, Metastasis staging system for lung cancer. Respirology. 2012;17(2):247-60.
58. Little AG, Gay EG, Gaspar LE, Stewart AK. National survey of non-small cell lung cancer in the United States: epidemiology, pathology and patterns of care. Lung Cancer. 2007;57(3):253-60.
59. Goldstraw P, Crowley J, Chansky K, Giroux DJ, Groome PA, Rami-Porta R, et al. The IASLC Lung Cancer Staging Project: proposals for the revision of the TNM stage groupings in the forthcoming (seventh) edition of the TNM Classification of malignant tumours. J Thorac Oncol. 2007;2(8):706-14.
60. Villalobos P, Wistuba, II. Lung Cancer Biomarkers. Hematol Oncol Clin North Am. 2017;31(1):13-29.
61. Minuti G, D'Incecco A, Cappuzzo F. Targeted therapy for NSCLC with driver mutations. Expert Opin Biol Ther. 2013;13(10):1401-12.
62. Dutch Guidelines regarding lung cancer. Adapted from www.oncoline.nl Access date 04-05-2017.
63. Vansteenkiste J, Crino L, Dooms C, Douillard JY, Faivre-Finn C, Lim E, et al. 2nd ESMO Consensus Conference on Lung Cancer: early-stage non-small-cell lung cancer consensus on diagnosis, treatment and follow-up. Ann Oncol. 2014;25(8):1462-74.
64. Scott WJ, Allen MS, Darling G, Meyers B, Decker PA, Putnam JB, et al. Video-assisted thoracic surgery versus open lobectomy for lung cancer: a secondary analysis of data from the American College of Surgeons Oncology Group Z0030 randomized clinical trial. J Thorac Cardiovasc Surg. 2010;139(4):976-81; discussion 81-3.
65. Group NM-aC, Arriagada R, Auperin A, Burdett S, Higgins JP, Johnson DH, et al. Adjuvant chemotherapy, with or without postoperative radiotherapy, in operable non-small-cell lung cancer: two meta-analyses of individual patient data. Lancet. 2010;375(9722):1267-77.
66. Solda F, Lodge M, Ashley S, Whitington A, Goldstraw P, Brada M. Stereotactic radiotherapy (SABR) for the treatment of primary non-small cell lung cancer; systematic review and comparison with a surgical cohort. Radiother Oncol. 2013;109(1):1-7.
67. van den Berg LL, Klinkenberg TJ, Groen HJ, Widder J. Patterns of Recurrence and Survival after Surgery or Stereotactic Radiotherapy for Early Stage NSCLC. J Thorac Oncol. 2015;10(5):826-31.

68. Auperin A, Le Pechoux C, Rolland E, Curran WJ, Furuse K, Fournel P, et al. Meta-analysis of concomitant versus sequential radiochemotherapy in locally advanced non-small-cell lung cancer. J Clin Oncol. 2010;28(13):2181-90.
69. Califano R, Abidin AZ, Peck R, Faivre-Finn C, Lorigan P. Management of small cell lung cancer: recent developments for optimal care. Drugs. 2012;72(4):471-90.
70. Pijls-Johannesma M, De Ruysscher D, Vansteenkiste J, Kester A, Rutten I, Lambin P. Timing of chest radiotherapy in patients with limited stage small cell lung cancer: a systematic review and meta-analysis of randomised controlled trials. Cancer Treat Rev. 2007;33(5):461-73.
71. Schneider BJ, Saxena A, Downey RJ. Surgery for early-stage small cell lung cancer. J Natl Compr Canc Netw. 2011;9(10):1132-9.
72. Auperin A, Arriagada R, Pignon JP, Le Pechoux C, Gregor A, Stephens RJ, et al. Prophylactic cranial irradiation for patients with small-cell lung cancer in complete remission. Prophylactic Cranial Irradiation Overview Collaborative Group. N Engl J Med. 1999;341(7):476-84.
73. Chansky K, Sculier JP, Crowley JJ, Giroux D, Van Meerbeeck J, Goldstraw P, et al. The International Association for the Study of Lung Cancer Staging Project: prognostic factors and pathologic TNM stage in surgically managed non-small cell lung cancer. J Thorac Oncol. 2009;4(7):792-801.
74. Sun Z, Aubry MC, Deschamps C, Marks RS, Okuno SH, Williams BA, et al. Histologic grade is an independent prognostic factor for survival in non-small cell lung cancer: an analysis of 5018 hospital- and 712 population-based cases. J Thorac Cardiovasc Surg. 2006;131(5):1014-20.
75. Sardari Nia P, Weyler J, Colpaert C, Vermeulen P, Van Marck E, Van Schil P. Prognostic value of smoking status in operated non-small cell lung cancer. Lung Cancer. 2005;47(3):351-9.
76. Kobayashi N, Toyooka S, Soh J, Ichimura K, Yanai H, Suehisa H, et al. Risk factors for recurrence and unfavorable prognosis in patients with stage I non-small cell lung cancer and a tumor diameter of 20 mm or less. J Thorac Oncol. 2007;2(9):808-12.
77. Zhou W, Heist RS, Liu G, Park S, Neuberg DS, Asomaning K, et al. Smoking cessation before diagnosis and survival in early stage non-small cell lung cancer patients. Lung Cancer. 2006;53(3):375-80.
78. Jamal A, King BA, Neff LJ, Whitmill J, Babb SD, Graffunder CM. Current Cigarette Smoking Among Adults - United States, 2005-2015. MMWR Morb Mortal Wkly Rep. 2016;65(44):1205-11.
79. Didkowska J, Wojciechowska U, Manczuk M, Lobaszewski J. Lung cancer epidemiology: contemporary and future challenges worldwide. Ann Transl Med. 2016;4(8):150.
80. Moolgavkar SH, Holford TR, Levy DT, Kong CY, Foy M, Clarke L, et al. Impact of reduced tobacco smoking on lung cancer mortality in the United States during 1975-2000. J Natl Cancer Inst. 2012;104(7):541-8.
81. Nagelhout GE, Levy DT, Blackman K, Currie L, Clancy L, Willemsen MC. The effect of tobacco control policies on smoking prevalence and smoking-attributable deaths. Findings from the Netherlands SimSmoke Tobacco Control Policy Simulation Model. Addiction. 2012;107(2):407-16.
82. Heckman CJ, Egleston BL, Hofmann MT. Efficacy of motivational interviewing for smoking cessation: a systematic review and meta-analysis. Tob Control. 2010;19(5):410-6.

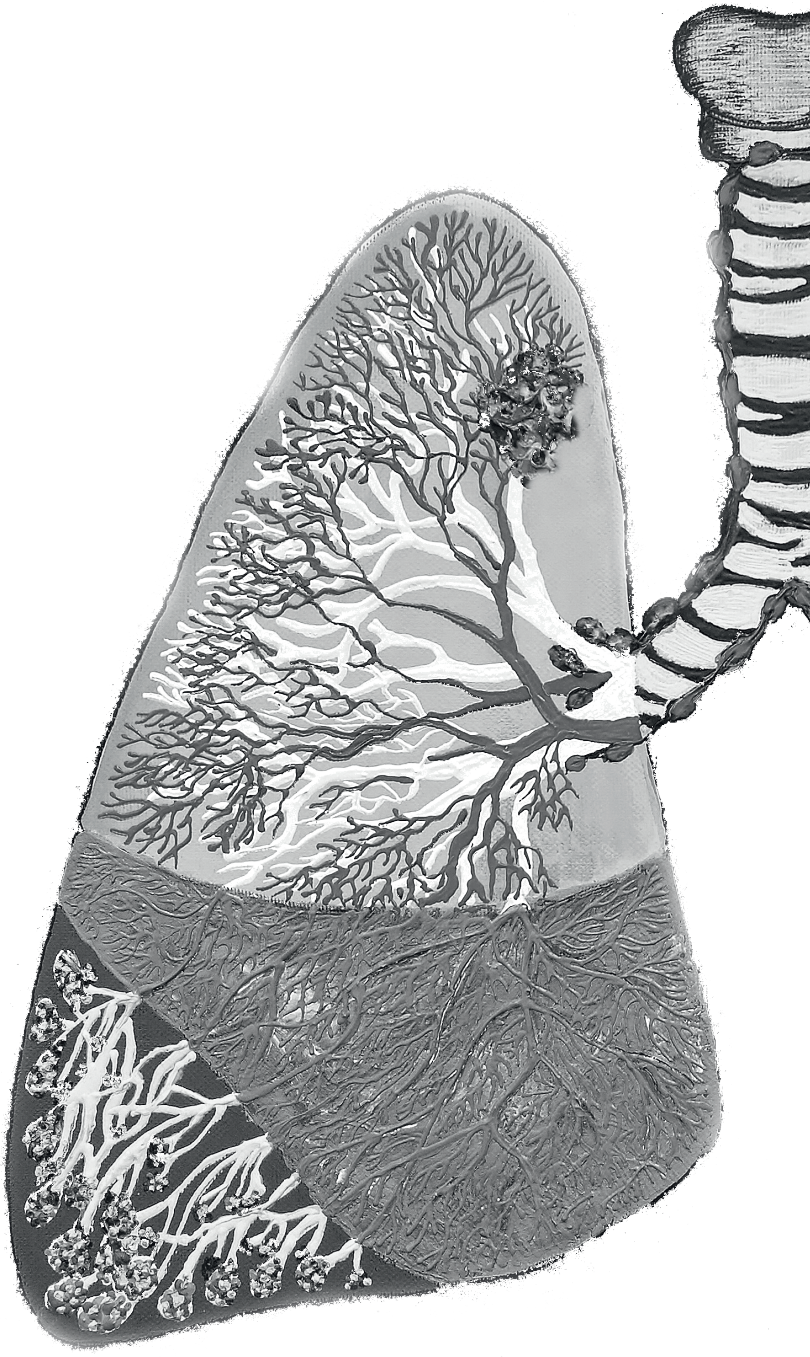
83. Cahill K, Lancaster T. Workplace interventions for smoking cessation. Cochrane Database Syst Rev. 2014(2):CD003440.
84. Yang P, Allen MS, Aubry MC, Wampfler JA, Marks RS, Edell ES, et al. Clinical features of 5,628 primary lung cancer patients: experience at Mayo Clinic from 1997 to 2003. Chest. 2005;128(1):452-62.
85. Mong C, Garon EB, Fuller C, Mahtabifard A, Mirocha J, Mosenifar Z, et al. High prevalence of lung cancer in a surgical cohort of lung cancer patients a decade after smoking cessation. J Cardiothorac Surg. 2011;6:19.
86. Manser R. Screening for lung cancer: a review. Curr Opin Pulm Med. 2004;10(4):266-71.
87. Manser RL, Irving LB, Byrnes G, Abramson MJ, Stone CA, Campbell DA. Screening for lung cancer: a systematic review and meta-analysis of controlled trials. Thorax. 2003;58(9):784-9.
88. Humphrey LL, Teutsch S, Johnson M, Force USPST. Lung cancer screening with sputum cytologic examination, chest radiography, and computed tomography: an update for the U.S. Preventive Services Task Force. Ann Intern Med. 2004;140(9):740-53.
89. Eddy DM. Screening for lung cancer. Ann Intern Med. 1989;111(3):232-7.
90. International Early Lung Cancer Action Program I, Henschke CI, Yankelevitz DF, Libby DM, Pasmantier MW, Smith JP, et al. Survival of patients with stage I lung cancer detected on CT screening. N Engl J Med. 2006;355(17):1763-71.
91. Sone S, Li F, Yang ZG, Honda T, Maruyama Y, Takashima S, et al. Results of three-year mass screening programme for lung cancer using mobile low-dose spiral computed tomography scanner. Br J Cancer. 2001;84(1):25-32.
92. Swensen SJ, Jett JR, Hartman TE, Midthun DE, Sloan JA, Sykes AM, et al. Lung cancer screening with CT: Mayo Clinic experience. Radiology. 2003;226(3):756-61.
93. 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.
94. Nawa T, Nakagawa T, Kusano S, Kawasaki Y, Sugawara Y, Nakata H. Lung cancer screening using low-dose spiral CT: results of baseline and 1-year follow-up studies. Chest. 2002;122(1):15-20.
95. Aberle Dr Fau - Berg CD, Berg Cd Fau - Black WC, Black Wc Fau - Church TR, Church Tr Fau - Fagerstrom RM, Fagerstrom Rm Fau - Galen B, Galen B Fau - Gareen IF, et al. The National Lung Screening Trial: overview and study design. (1527-1315 (Electronic)).
96. National Lung Screening Trial Research T, 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.
97. Patz EF, Jr., Greco E, Gatsonis C, Pinsky P, Kramer BS, Aberle DR. Lung cancer incidence and mortality in National Lung Screening Trial participants who underwent low-dose CT prevalence screening: a retrospective cohort analysis of a randomised, multicentre, diagnostic screening trial. Lancet Oncol.May;17(5)(1474-5488 (Electronic)):590-9.
98. 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.

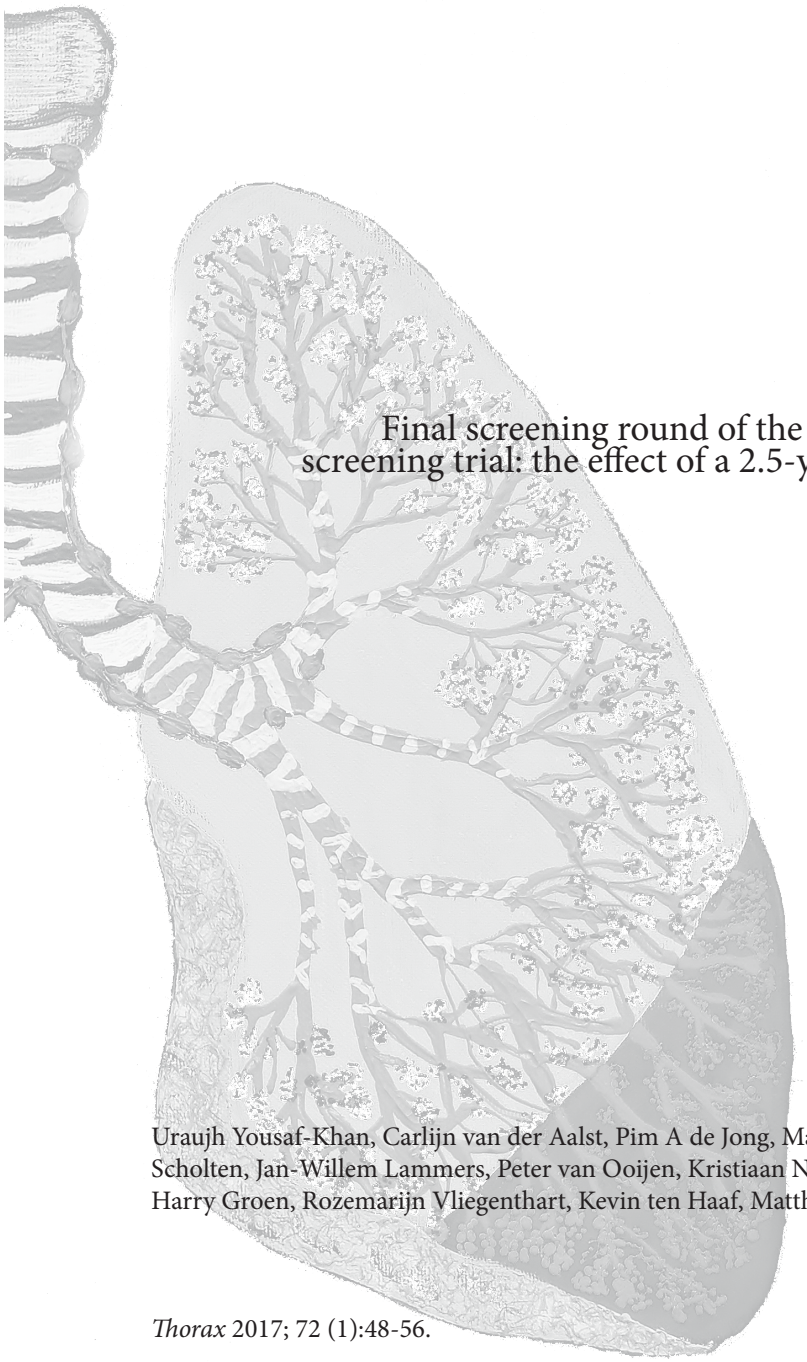
99. Humphrey LL, Deffebach M, Pappas M, Baumann C, Artis K, Mitchell JP, et al. Screening for lung cancer with low-dose computed tomography: a systematic review to update the US Preventive services task force recommendation. Ann Intern Med. 2013;159(6):411-20.
100. van Iersel CA, de Koning HJ, Draisma G, Mali WP, 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.
101. 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. J Thorac Oncol. 2009;4(5):608-14.
102. 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. Eur J Cancer Prev. 2012;21(3):308-15.
103. Baldwin DR, Duffy SW, Wald NJ, Page R, Hansell DM, Field JK. UK Lung Screen (UKLS) nodule management protocol: modelling of a single screen randomised controlled trial of low-dose CT screening for lung cancer. Thorax. 2011;66(4):308-13.
104. 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.
105. 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.
106. 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.
107. 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 Med. 2009;361(23):2221-9.
108. 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.
109. Wille MM, Dirksen A, Ashraf H, Saghir Z, Bach KS, Brodersen J, et al. Results of the Randomized Danish Lung Cancer Screening Trial with Focus on High-risk Profiling. (1535-4970 (Electronic)).
110. Paci E, Puliti D, Lopes Pegna A, Carrozzi L, Picozzi G, Falaschi F, et al. Mortality, survival and incidence rates in the ITALUNG randomised lung cancer screening trial. Thorax. 2017.
111. Infante M, Cavuto S, Lutman FR, Passera E, Chiarenza M, Chiesa G, et al. Long-Term Follow-up Results of the DANTE Trial, a Randomized Study of Lung Cancer Screening with Spiral Computed Tomography. Am J Respir Crit Care Med. 2015;191(10):1166-75.
112. Aberle DR, DeMello S Fau - Berg CD, Berg Cd Fau - Black WC, Black Wc Fau - Brewer B, Brewer B Fau - Church TR, Church Tr Fau - Clingan KL, et al. Results of the two incidence screenings in the National Lung Screening Trial. (1533-4406 (Electronic)).

113. 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.
114. 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.
115. Horeweg N, Scholten ET, de Jong PA, van der Aalst CM, Weenink C, Lammers JW, et al. Detection of lung cancer through low-dose CT screening (NELSON): a prespecified analysis of screening test performance and interval cancers. Lancet Oncol. 2014;15(12):1342-50.
116. 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.
117. Ten Haaf K, de Koning HJ. Overdiagnosis in lung cancer screening: why modelling is essential. J Epidemiol Community Health. 2015;69(11):1035-9.
118. van der Aalst CM, van Iersel CA, van Klaveren RJ, Frenken FJ, 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.
119. Xu DM, Gietema H, de Koning H, Vernhout R, Nackaerts K, Prokop M, et al. Nodule management protocol of the NELSON randomised lung cancer screening trial. Lung Cancer. 2006;54(2):177-84.
120. Xie X, Zhao Y Fau - Snijder RA, Snijder Ra Fau - van Ooijen PMA, van Ooijen Pm Fau - de Jong PA, de Jong Pa Fau - Oudkerk M, Oudkerk M Fau - de Bock GH, et al. Sensitivity and accuracy of volumetry of pulmonary nodules on low-dose 16- and 64-row multi-detector CT: an anthropomorphic phantom study. (1432-1084 (Electronic)).
121. Wang Y, van Klaveren RJ, van der Zaag-Loonen HJ, de Bock GH, Gietema HA, Xu DM, et al. Effect of nodule characteristics on variability of semiautomated volume measurements in pulmonary nodules detected in a lung cancer screening program. Radiology. 2008;248(2):625-31.
122. Horeweg N, van Klaveren RJ, Groen HJ, Lammers JW, Weenink C, Nackaerts K, et al. Blinded and uniform cause of death verification in a lung cancer CT screening trial. Lung Cancer. 2012;77(3):522-5.

Part I

the optimization of the NELSON screening rounds





CHAPTER II

Final screening round of the NELSON lung cancer screening trial: the effect of a 2.5-year screening interval

Uraujh Yousaf-Khan, Carlijn van der Aalst, Pim A de Jong, Marjolein Heuvelmans, Ernst Scholten, Jan-Willem Lammers, Peter van Ooijen, Kristiaan Nackaerts, Carla Weenink, Harry Groen, Rozemarijn Vliegenthart, Kevin ten Haaf, Matthijs Oudkerk, Harry de Koning

Thorax 2017; 72 (1):48-56.

ABSTRACT

Background
In the USA annual lung cancer screening is recommended. However, the optimal screening strategy (eg, screening interval, screening rounds) is unknown. This study provides results of the fourth screening round after a 2.5-year interval in the Dutch-Belgian Lung Cancer Screening trial (NELSON).

Methods
Europe's largest, sufficiently powered randomised lung cancer screening trial was designed to determine whether low-dose CT screening reduces lung cancer mortality by ≥25% compared with no screening after 10 years of follow-up. The screening arm (n=7,915) received screening at baseline, after 1 year, 2 years and 2.5 years. Performance of the NELSON screening strategy in the final fourth round was evaluated. Comparisons were made between lung cancers detected in the first three rounds, in the final round and during the 2.5-year interval.

Results
In round 4, 46 cancers were screen-detected and there were 28 interval cancers between the third and fourth screenings. Compared with the second round screening (1-year interval), in round 4 a higher proportion of stage IIIb/IV cancers (17.3% vs 6.8%, p=0.02) and higher proportions of squamous-cell, bronchoalveolar and small-cell carcinomas (p=0.001) were detected. Compared with a 2-year interval, the 2.5-year interval showed a higher non-significant stage distribution (stage IIIb/IV 17.3% vs 5.2%, p=0.10). Additionally, more interval cancers manifested in the 2.5-year interval than in the intervals of previous rounds (28 vs 5 and 28 vs 19).

Conclusions
A 2.5-year interval reduced the effect of screening: the interval cancer rate was higher compared with the 1-year and 2-year intervals, and proportion of advanced disease stage in the final round was higher compared with the previous rounds.

2.1. INTRODUCTION

Lung cancer remains the leading cause of cancer death worldwide, mainly due to its advanced stage at the time of diagnosis [1]. Based on the results of the National Lung Screening Trial (NLST), the US Preventive Services Task Force recommends annual lung cancer screening with CT [2, 3]. People eligible for screening are aged 55 years through 80 years, have smoked at least 30 pack-years, and currently smoke or have quit within the past 15 years [4, 5]. However, little is known about the effect of longer screening intervals in lung cancer screening trials: thus far, only the Multicentric Italian Lung Detection trial which consisted of two low-dose CT (LDCT) arms (annual vs biennial screening), reported no differences in mortality or in screening test performances between the two arms (n=1,190 and n=1,186) [6, 7].

The Dutch-Belgian Lung Cancer Screening trial (NELSON) is the largest European randomized lung cancer screening trial, which was designed to investigate whether LDCT screening reduces lung cancer mortality by ≥25% compared with no screening after 10 years of follow-up [8, 9]. The trial randomised (1:1) 15,822 current or former smokers into a screening group and a control group. Compared with the NLST control group who received screening by chest radiography, NELSON control group participants received no screening. Furthermore, the NELSON screening group received LDCT screening at baseline (round 1), after 1 year (round 2), after 3 years (round 3) and after 5.5 years after baseline (round 4), whereas the NLST provided three annual screenings [10]. The use of variable screening intervals in one LCDT arm in the sufficiently powered NELSON trial is unique and presents an opportunity to investigate the influence of the intervals on the screening test performances (eg, lung cancer detection rate, false-positive (FP) rate) and the characteristics of screening-detected lung cancers.

Analyses of the first three rounds of the NELSON trial indicated that a 2-year interval between the second and the third screening rounds did not lead to a significantly higher proportion of advanced stage lung cancers compared with a 1-year screening interval between the first and second rounds [11]. Furthermore, the lung cancer detection rate was relatively stable across the first three rounds [11-13]. Analyses also indicated that, despite the 2-year interval between the second and third rounds, specificity and sensitivity of the first three rounds were higher compared with other screening trials, which suggests that lung cancer screening using biennial screening regimens after an initial screening round could be effective [14].

The primary aim of this study is to investigate the additional value of the final fourth screening round, 2.5 years after the previous screening round. The performance of the NELSON screening strategy in the final screening round is evaluated, and comparisons are made between the lung cancers detected in the first three rounds, those detected in the final round and cancers detected in the 2.5-year screening interval between the third and fourth rounds (ie, interval cancers).

2.2. METHODS

NELSON trial

Details of the design and conduct of the NELSON trial have been reported previously [8, 9]. In brief, eligible participants were selected after completing questionnaires about general health, lifestyle and smoking habits. Based on this information, persons aged 50–75 years, who had smoked ≥15 cigarettes per day for ≥25 years or ≥10 cigarettes per day for ≥30 years, and who were current smokers or former smokers with cessation ≤10 years ago, were invited to participate in the NELSON trial. Eventually, 15 822 eligible high-risk subjects for developing lung cancer participated in this population-based randomised trial. The primary aim of NELSON is to determine whether LDCT screening reduces lung cancer mortality by ≥25% compared with no screening after 10 years of follow-up [8] .

To perform a fourth screening round an additional informed consent was obtained, as the original protocol consisted of only three screening rounds. The final screening round was conducted from November 2009 through March 2012.

Study population

For this study, all 7,915 participants randomised to the screening arm were included.

Screening procedures

Screening group participants were invited to one of the four screening sites (University Medical Centre Groningen, University Medical Centre Utrecht and Kennemer Gasthuis Haarlem in the Netherlands, and University Hospital Gasthuisberg Leuven in Belgium). For the screening, 16-detector or, in later rounds 64-detector CT scanners in low-dose setting were used, without the administration of intravenous contrast media. Images were analysed using semiautomated software (LungCARE, version Somaris/5 VA70C-W, Siemens Medical Solutions) [9, 15]. The analysis included the semiautomated segmentation of nodules and determination of the nodule volume. In case the software was not able to segment a nodule accurately, the diameter was measured manually by the radiologist [16]. In the first two rounds, two radiologists independently reviewed the images. In case of a discrepancy, a third expert reader made the final decision. In the last two rounds, a single reading was performed by a radiologist with at least 6 years of experience in thoracic imaging. Wang et al [17] showed that there was no benefit for double reading consensus with the use of semiautomated software. More detailed descriptions of the equipment and execution of the screening examination have been provided in previous reports [9, 15, 18]

Screening outcomes and the nodule management protocol

The screening test had three possible results, depending on the presence of nodules, nodule volume and volume doubling time (VDT): negative, indeterminate or positive [10]. Negative results led to invitation to the next screening round, or in case of the final round to the end of the screening programme. Indeterminate results led to invitation for a repeat scan (after 6–8 weeks or after 12 months, depending on nodule size and screening round) in order to classify the final result as positive or negative, based on volume change (growth) and growth rate, expressed in VDT [10]. Positive results led to referral to a pulmonologist for a diagnostic workup. If lung cancer was diagnosed, a participant received treatment according to (inter)national guidelines. Medical data of these participants were collected prospectively. If a workup after a positive screening did not lead to lung cancer diagnosis, participants were invited for the next screening round, or, in case of the fourth round, to the end of the screening programme.

Nodule management protocol

Briefly, in case of newly detected solid nodules and the solid component of part-solid nodules, the volume determined the screening result: <50 mm3 was negative, 50–500 mm3 was indeterminate and >500 mm3 was positive [10, 12, 15, 17, 18]. In case of previously detected nodules, evaluation was based on growth (defined as change in volume) and VDT. If volume growth was <25%, the screening result was negative, and if the volume growth was ≥25%, the VDT of the nodule was calculated: for nodules with VDT of 400–600 days, the result was indeterminate, and the result was positive if the VDT was <400 days and/or a new solid component emerged in a previously non-solid nodule.

Definitions

A regular round scan is the first CT examination performed for a specific participant in one of the predefined screening rounds. Follow-up scans are repeat scans which were performed in between screening rounds if a participant had an indeterminate result in one of the four regular scans. The result of a regular scan was defined as the result of the first CT examination in a screening round, while the definitive outcome of the screening round was determined after inclusion of the results of the repeat scans performed within that particular screening round.

Lung cancers diagnosed by a pulmonologist within 24 months after referral for a positive screening were defined as screening-detected lung cancers. Interval lung cancers were defined as: lung cancers diagnosed after a negative screening test and before a next screening round; lung cancers diagnosed after an indeterminate screening test, without a follow-up CT scan before the next screening round; or lung cancers diagnosed after a positive screening result if the diagnostic workup initiated for the positive screening result did not yield a diagnosis of lung cancer, and the diagnosis was made later because symptoms had triggered diagnostic assessment that eventually yielded diagnosis of lung cancer. Overall lung cancer detection

rate was defined as the number of screening-detected lung cancers divided by the number of screened participants. Lung cancer detection rate of a round was defined as the number of screening-detected lung cancers in that round divided by the number of screened participants in that round. An FP test result was defined as a positive result in a participant when lung cancer was not diagnosed after referral to a pulmonologist for a diagnostic workup; a true-positive (TP) test result was defined as a positive result in a participant diagnosed with lung cancer after workup by a pulmonologist. The overall FP rate result of a screening round was defined as the total number of FP screenings divided by the total number of scans performed in that round; the overall FP rate of the NELSON trial was defined as the total number of FP screenings across all screening rounds divided by the total numbers of scans performed across the four screening rounds.

Statistical analyses

Continuous variables were tested for normality by using the Kolmogorov-Smirnov test and examining Q-Q-plots. None of the tested variables were distributed normally, so they were described by using medians and IQRs. Differences between nominal variables were calculated by using a $\chi 2$ test, and differences between categorical variables by using a Mann-Whitney U test. To calculate the 95% CI for the lung cancer detection rate, positive predictive value, the TP rate, and the FP rate, bootstrapping was performed with 5000 samples. For all analyses, a p value <0.05 was considered statistically significant, and IBM SPSS Statistics V.21 was used for all analysis.

2.3. RESULTS

The fourth screening round

The participation rates in the first three rounds were: 7,557 (95.5%) in round 1, 7,295 (92.2%) in round 2 and 6,922 (87.5%) in round 3. All eligible participants from the third round were invited to participate in the final round; in total 6,735 of the 6,922 participants screened in the third screening round were eligible, since they were alive and had not been diagnosed with lung cancer. Of these eligible participants, 80.7% (5,437/6,735) responded positively, and 97.1% (5,279/5,437) of these attended the final round (**Figure 1**). Participants with solely negative screening results were more willing to participate in the final round, compared with those with at least one non-negative screening result in the previous screening rounds (p=0.006, data not shown). Moreover, a higher proportion of current smokers attended the last round compared with former smokers (54.5% vs 45.5%, p=0.04).
Figure 2 presents an overview of the screening outcomes of the final round: 5,380 scans were performed of which 98.1% (5,279/5,380) were regular round scans, and 1.9% (101/5,380) were follow-up scans performed to assess the VDT of intermediate nodules.
In total 5,336 nodules were detected on 2,507 of 5,380 performed scans in this round. Of these, 35.2% were indeterminate (NODCAT III), and 2.8% were potentially malignant (NODCAT IV). Most of the nodules were solid (93.9%) and had a VDT of >600 days (98.1%; **Table 1**).

In round 4, 46 lung cancers were detected in 43 of 105 participants with a positive result, representing a TP rate of 41.0% (43/105). Of these cancers, 60.9% were detected in stage I, 15.2% in stage II, 10.8% in stage III and 13.1% in stage IV. Most cancers were localised in the right lung (67.4%, online supplementary **Table S1**). Four participants with lung cancer had symptoms before diagnosis. Half of the lung cancers were adenocarcinomas, and these tended to be more frequently detected in lower stage (stage I–IIIa) compared with small-cell lung cancers which were diagnosed in stage IIIa or stage IV (small cell lung carcinoma (SCLC); p=0.06). No significant correlations were found between disease stage and the following factors: age (p=0.81), gender (p=0.38), smoking status (p=0.89) or starting age of smoking (p=0.28; data not shown) (**Table 2**).

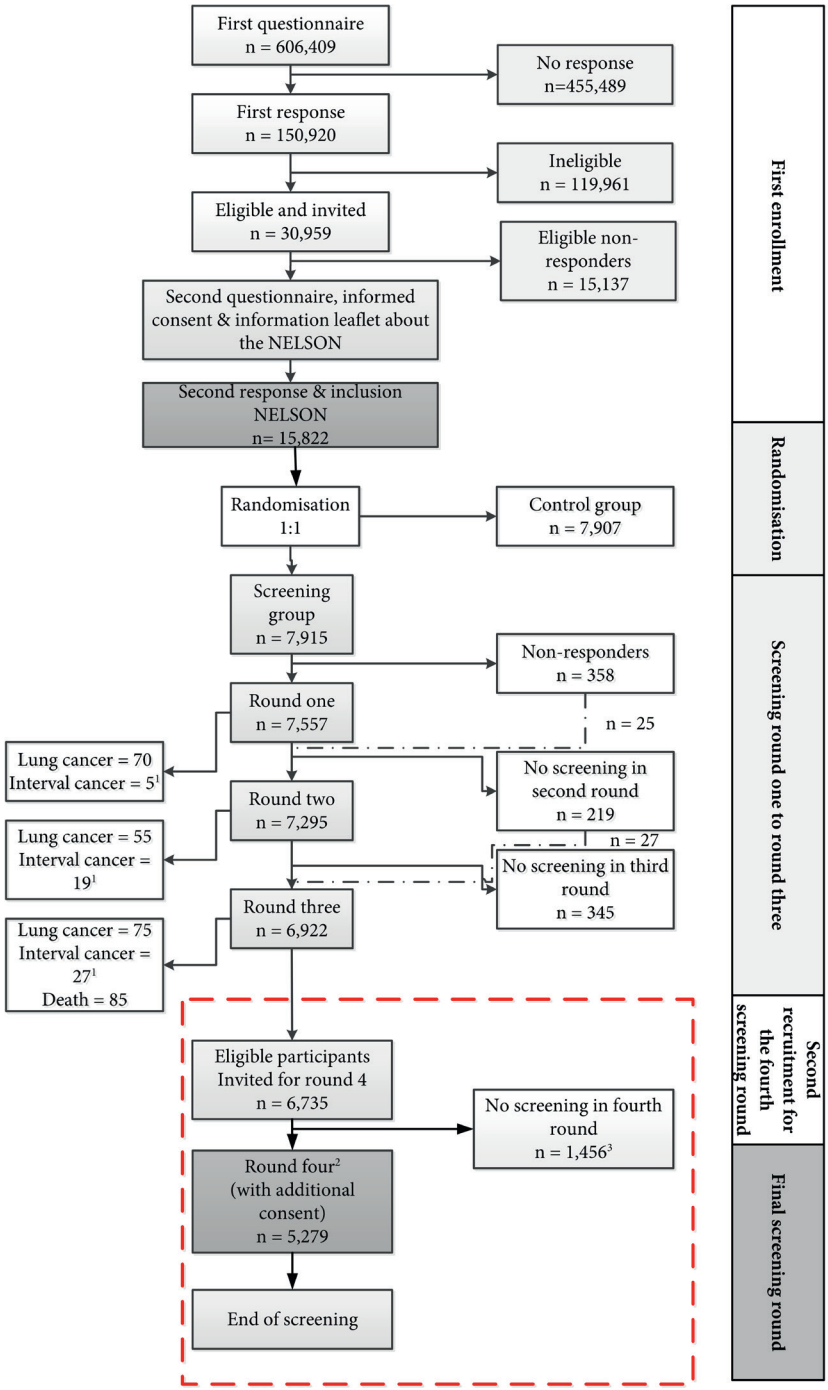

First questionnaire
n = 606,409
No response
n=455,489
First response
n = 150,920
Ineligible
n = 119,961
Eligible and invited
n = 30,959
Eligible non-responders
n = 15,137
Second questionnaire, informed consent & information leaflet about the NELSON
Second response & inclusion NELSON
n= 15,822
Randomisation
1:1
Control group
n = 7,907
Screening group
n = 7,915
Non-responders
n = 358
Round one
n = 7,557
n = 25
Lung cancer = 70
Interval cancer = 5[1]
No screening in second round
n = 219
Round two
n = 7,295
n = 27
Lung cancer = 55
Interval cancer = 19[1]
No screening in third round
n = 345
Round three
n = 6,922
Lung cancer = 75
Interval cancer = 27[1]
Death = 85
Eligible participants
Invited for round 4
n = 6,735
No screening in fourth round
n = 1,456[3]
Round four[2]
(with additional consent)
n = 5,279
End of screening
First enrollment
Randomisation
Screening round one to round three
Second recruitment for the fourth screening round
Final screening round

Figure 1: flow chart of the NELSON lung cancer screening study. Dashed lines:
25 participants did not receive screening in the first screening round but were screened in the second screening round; 27 participants received no screening in the second screening round, but were screened in the third round. Red dashed box: screened participants invited for the fourth screening round.

[1]Interval cancer data of only Dutch screening group participants. [2]Only participants who gave their additional consent were screened in the last screening round. [3]Reasons for no further screening: 49 participants weren't traceable, 40 participants declined due to illness, 4 participants found the participating centre too far, 3 participants didn't receive travelling expenses, 4 participants had a negative experience with other trials, 3 participants declined because of a ill family relative and 1,353 participants didn't respond. Eventually, 5,279 participants provided additional consent and were screened in the final screening round.

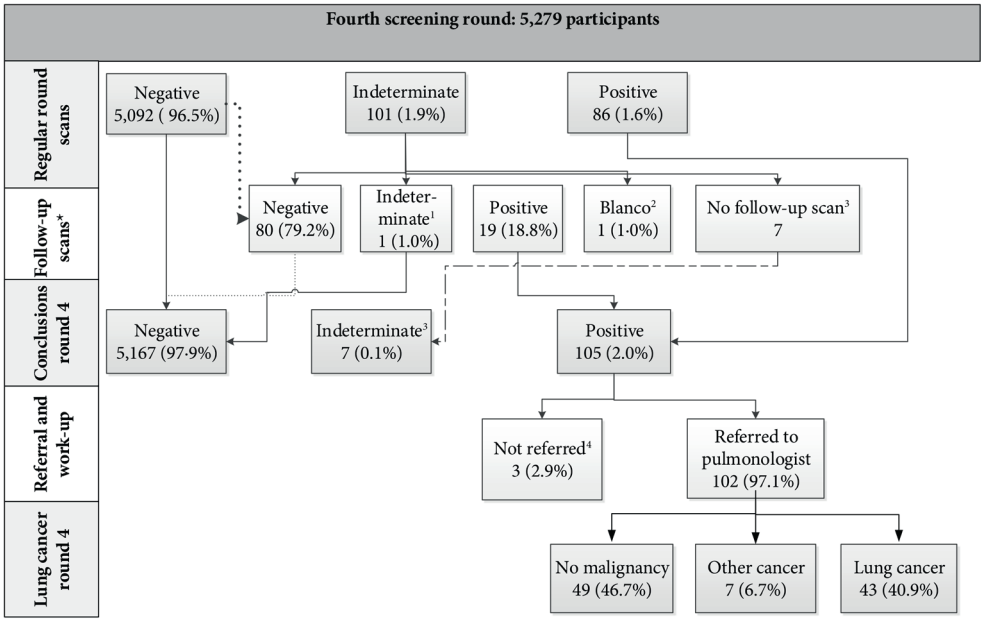


Figure 2: screening results of the fourth screening round

* Follow-up scans are scans performed after an indeterminate screening result.
Dotted line: four participants with a negative regular scan result received accidently a follow-up scan. Their screening results remained negative.

[1] In the follow-up scan one participant received an indeterminate result, so the participant received a second follow-up scan: the final scan was negative.

[2] From one participant the scan data was lost, the participant received a new follow-up scan. The final screening result was negative.

[3] Seven participants received no follow-up scan (dashed line): two participants declined further screening, one participant did not respond anymore and four participants had an indeterminate screening result but were accidentally not invited for a follow-up scan. Hence, for these seven participants the final screening outcome remained indeterminate. As these participants received no follow-up scans, the sum of follow-up scans is: 80 + 1+ 19 + 1 = 101.

[4] Three participants with a positive screening were not referred to a pulmonologist: in one case the radiologist judged the growing nodule, despite a volume doubling time of less than 400 days, as non-malignant; one participant was already diagnosed with an interval cancer and should not have been invited and screened in the final round; and in one case the work-up was started, but ended shortly after the patient was deemed too ill to undergo invasive diagnostics.

Table 1: overview of the nodules detected in the scans performed in the fourth screening round

		n	%
Total scans performed in round four		5,380	100.0
Scans with nodules		2,507	46.6
Scans without nodules		2,873	53.4
Total nodules detected on scans performed in round 4		5336	100.0
Nodule size category†			
NODCAT I*	Benign	254	4.8
NODCAT II**	Non-significant small	3,057	57.3
NODCAT III***	Indeterminate	1,875	35.2
NODCAT IV****	Potentially malignant	149	2.8
Nodule growth category†			
GROWCAT A‡		2,348	98.1
GROWCAT B‡‡		14	0.6
GROWCAT C‡‡‡		29	1.3
Type			
Solid		5,012	93.9
Partial solid		39	0.7
Non-solid		47	0.9
Disappeared nodules		161	3.0
Calcified/benign		97	1.6
Location			
Right lung§		3,109	58.3
Left lung		2,227	41.7

*NODCAT I: Nodule with benign characteristics such as benign calcification patterns or fat depositions
**NODCAT II: Solid nodules with volume < 50 mm^3; Pleural-based solid nodules with minimum diameter < 5 mm; Non solid component part solid SSNs with average diameter < 8 mm; Solid component part solid SSNs with volume < 50 mm^3; Non solid nodules with average diameter < 8 mm
***NODCAT III: Solid nodules with volume 50 - 500 mm^3; Pleural-based solid nodules with minimum diameter 5 - 10 mm; Non solid component part solid SSNs with average diameter ≥ 8 mm; Solid component part solid SSNs with volume 50 - 500 mm^3; Non solid nodules with average diameter ≥ 8 mm
****NODCAT IV: Pleural-based solid nodules with minimum diameter > 10 mm; Solid component part solid SSNs with volume > 500 mm^3
† One nodule without outcome
‡GROWCAT A: Percentage volume change ≥ 25% and VDT > 600 days
‡‡GROWCAT B: Percentage volume change ≥ 25% and VDT 400 - 600 days
‡‡‡GROWCAT C: Percentage volume change ≥ 25% and VDT < 400 days; New solid component in previously non solid nodule
§One nodule was located in the trachea
GROWCAT, nodule growth category; NODCAT, nodule size category; VDT, volume doubling time.

Table 2: clinical features of screening-detected lung cancers in the fourth screening round

	Cancer stage							
	Ia	**Ib**	**IIa**	**IIb**	**IIIa**	**IIIb**	**IV**	**Total histology**
	n, %	n, %	n, %	n, %	n, %	n, %	n, %	n, %
Histology								
Adenocarcinoma	11* (47.8)	3 (13.0)	2 (8.7)	3 (13.0)	2 (8.7)	1 (4.3)	1 (4.3)	23 (50.0)†
Squamous-cell carcinoma	5 (50.0)	2 (20.0)	1 (10.0)	1 (10.0)	-	-	1 (10.0)	10 (21.7)
BAC	4‡ (75.0)	-	-	-	-	-	-	4 (8.7)
Small-cell carcinoma	-	-	-	-	1 (33.3)	-	2 (66.7)	3 (6.5)
NSCLC NOS	1 (25.0)	-	-	-	-	1 (25.0)	2 (50.0)	4 (8.7)
No diagnosis possible	1 (50.0)	1 (50.0)	-	-	-	-	-	2 (4.4)
Total Cancer stage n, %	22 (47.8)§**	6 (13.1)	3 (6.5)	4 (8.7)	3 (6.5)	2 (4.3)	6 (13.1)	46 (100)

*For example 11 (47,8) should be read as: 11 adenocarcinoma were detected in stage Ia, which corresponds with a 47.8% (11/23).
†For example 23 (50.0) should be read as: 23 adenocarcinomas were detected in the fourth screening round, this corresponds with a 50,0% of all the screening-detected lung cancers in this round.
‡Of which one was a pTisN0 non-mucinous adenocarcinoma in situ. This cancer was grouped together with a stage Ia lung cancer.
§For example 22 (47.8) should be read as: 22 cancers were detected in stage 1a, which corresponds with a 47.8% cancers detected in stage Ia (22/46).
BAC, bronchoalveolar carcinoma; NSCLC NOS, non-small cell lung carcinoma not otherwise specified.

Screening outcomes of 2.5-year versus 1-year screening intervals

Compared with round 2, performed after a 1-year screening interval, a lower proportion of stage I (60.9% vs 75.9%) and a higher proportion of stage IIIb/IV (17.3% vs 6.8%) cancers was detected in the final round (p=0.02, **Table 3**). Relative to the results of a 1-year screening interval, higher proportions of squamous-cell carcinomas (SQM), bronchoalveolar carcinomas (BAC) and SCLC were detected (p=0.001, **Table 4**). In round 4, no large cell carcinomas, large cell neuroendocrine carcinomas (LCNECs) or carcinoids were detected. The locations of lung cancer or proportion of female participants with lung cancer did not differ between the second and fourth rounds (p=0.91 and p=0.78, respectively, online supplementary **Table S1**).

Screening outcomes of 2.5-year versus 2-year screening intervals

In the final round (performed after an interval of 2.5 years) a lower proportion of cancers was diagnosed in stage I (60.9% vs 72.7%) and a higher proportion in stage IIIb/IV (17.3% vs 5.2%), compared with round 3 with a 2-year screening interval. However, this difference did not reach statistical significance (p=0.10). Compared with the final round, in round 3 more cancers of other histology types (two large-cell carcinomas, two carcinoids, one LCNEC, one mixed non-small cell lung carcinoma (NSCLC)/SCLC and seven without a diagnosis) were detected (p=0.06). The localisations of lung cancers or proportion of female participants with lung cancer did not differ between the third and the last screening rounds (p=0.66 and p=0.73, respectively, online supplementary **Table S1**).

Screening test performance across the four screening rounds

The lung cancer detection rate in the fourth round was slightly lower compared with the detection rate in the third round, however, not statistically different (0.8%, 95% CI (0.6% to 1.1%) vs 1.1%, 95% CI (0.8% to 1.3%)). Compared with the first and second rounds, no differences in lung cancer detection rates were observed (**Table 5**). The ratio of the TP and FP results tended to improve over time, from 0.69 in round 1 to 0.72 in round 2 and 0.83 in round 3. However, in the last screening round it dropped to 0.69. The other screening test performances did not differ.

Table 3: stage distribution of screening-detected lung cancers of all rounds

	Round 1			p-value*	Round 2			p-value†	Round 3			p-value‡	Round 4		
	n	**%**	**Cum %**	0.37	**n**	**%**	**Cum %**	0.02	**n**	**%**	**Cum %**	0.10	**n**	**%**	**Cum %**
Stage															
Ia	44	59.5	59.5		43	74.1	74.1		50	64.9	64.9		22	47.8	47.8
Ib	4	5.4	64.9		1	1.7	75.9		6	7.8	72.7		6	13.0	60.9
IIa	7	9.5	74.3		4	6.9	82.8		-	-	76.6		3	6.5	67.4
IIb	-	-	-			-	-		3	3.9	-		4	8.7	76.1
IIIa	10	13.5	87.8		6	10.3	93.1		14	18.2	94.8		3	6.5	82.6
IIIb	4	5.4	93.2		2	3.4	96.6		1	1.3	96.1		2	4.3	87.0
IV	5	6.8	100		2	3.4	100		3	3.9	100		6	13.0	100
Total	74	100	-		58	100	-		77	100	-		46	100	-

*p-value: comparison of stage distribution of the screening-detected lung cancers of round 1 vs. round 4.
†p-value: comparison of stage distribution of the screening-detected lung cancers of round 2 vs. round 4.
‡ p-value: comparison of stage distribution of the screening-detected lung cancers of round 3 vs. round 4.
CUM %, cumulative percentage.

Table 4: histology of screening-detected lung cancers of all rounds

	Round 1		p-value*	Round 2		p-value†	Round 3		p-value‡	Round 4	
	n	%	0.01	n	%	0.001	n	%	0.06	n	%
Histology											
Adenocarcinoma	35	47.3		35	60.3		37	48.1		23	50.0
Squamous cell carcinoma	15	20.3		3	5.2		16	20.8		10	21.7
BAC[1]	2	2.7		3	5.2		6	7.8		4	8.7
SCLC[2]	1	1.4		2	3.4		5	6.5		3	6.5
NSCLC, not specified	2	2.7		-	-		-	-		4	8.7
Others	19§	25.7		15¶	25.9		13**	16.9		2††	4.3
Total	74	100		58	100		77	100		46	100

*p-value: difference in histology between the first and last screening round. Fisher's Exact test was used.
†p-value: difference in histology between the second and last screening round. Fisher's Exact test was used.
‡p-value: difference in histology between the third and last screening round. Fisher's Exact test was used.
§Others in round 1: 5 large cell carcinomas, 1 mixed NSCLC/SCLC, 4 carcinoids, 1 others, 3 no diagnosis was possible, 2 adenosquamous and 3 large cell neuro-endocrine carcinoma
¶Others in round 2: 10 large-cell carcinomas, 2 adenosquamous cell carcinoma and 3 cancers without a diagnosis were included in the group of others; in round 4: 2 cancers without a possible diagnosis were included in the group of others.
**Others in round 3: 2 large-cell carcinomas, 1 mixed NSCLC/SCLC, 2 carcinoids, 1 large-cell neuro-endocrine carcinoma and 7 without a diagnosis; in round 4: 2 cancers without a diagnosis.
††Others in round 4: 2 with no diagnosis possible. No other types of lung cancers were detected in this round.
BAC,bronchoalveolar carcinoma; SCLC, small-cell carcinoma.

Table 5: screening test performance across the four screening rounds

	R1*	95% CI	R2*	95% CI	R3*	95% CI	R4	95% CI
Lung cancer detection rate, %	0.9	0.7-1.2	0.8	0.6-1.0	1.1	0.8-1.3	0.8	0.6 - 1.1
Positive predicted value, %	35.5	28.4-42.1	42.0	34.4-49.6	45.5	37.6-53.5	41.0	31.6 - 50.5
False positive (FP) rate after positive screening, %	64.5	57.9-71.6	58.0	50.4-65.6	54.5	46.7-62.4	59.0	49.5 - 68.4
Ratio TP/FP	0.69	-	0.72	-	0.83	-	0.69	-
Overall FP rate†	-	-	-	-	-	-	1.2	-
Number needed to screen to detect 1 lung cancer	108	-	133	-	92	-	123	-

* Screening test performances across the first three rounds. [12]

† This is the overall FP rate of the NELSON trial across all four screening rounds.

TP, true positive.

Interval cancers diagnosed between the third and fourth screening rounds

Participants with an interval cancer diagnosed between the third and fourth screening rounds were slightly older (p=0.06) and had smoked more pack-years (<0.001) compared with participants screened in the fourth round (**Table 6**). Relative to the previous rounds, in this 2.5-year interval a higher proportion of participants was diagnosed with an interval cancer (**Figure 3**).

In the first 24 months after the third round, 12 participants were diagnosed with an interval cancer, while in the last 6 months 16 extra participants were diagnosed with an interval cancer. The median age of participants with an interval cancer in the first 24 months was slightly lower than that of the participants with an interval cancer in the last 6 months (64.2 vs 65.4, respectively, p=0.05). No differences were seen in stage distribution (p=0.77), histology (p=0.32), proportion of women (p=0.29) or in proportion of current smokers (p=0.22) between the interval cancers diagnosed in the first 2 years after the third round or in the last 6 months before the fourth screening round (see online supplementary **Tables S2–4**).

Compared with the screening-detected cancers in the final round, the interval cancers between the third and fourth screening rounds were more often diagnosed at stage IIIb/IV (64.3% vs 17.3%, p<0.001). They were also more often SCLC (10.7% vs 6.5%), large cell-carcinoma and LCNEC (14.3% vs 0%), and less often adenocarcinoma (32.1% vs 50.0) or BAC (0% vs 8.7%, p=0.02; all data not shown) compared with the lung cancers detected in round 4.

Table 6: baseline characteristics of participants

Characteristics	Participants with interval cancer in 2.5-year interval after R3 (n=27)*	Participants screened in R4 (n = 5,279)	Participants with screening-detected lung cancer in R4 (n=43)	p-value
Male n, (%)	23 (85.2)	4,437 (84.1)	37 (86.0)	0,93
Age median, (IQR)	59,0 (8.0)	58,0 (8.0)	60,0 (9.0)	0.06
Current smokers n, (%)	18 (66.7)	2,878 (54.5)	27 (62.8)	0.25
Pack years median, (IQR)	49,5 (28.8)	38,0 (19.8)	43,7 (23.0)	<0.001

*27 participants with 28 interval cancer in the 2.5-year interval; p-value: across all subgroups, Kruskal-Wallis test.

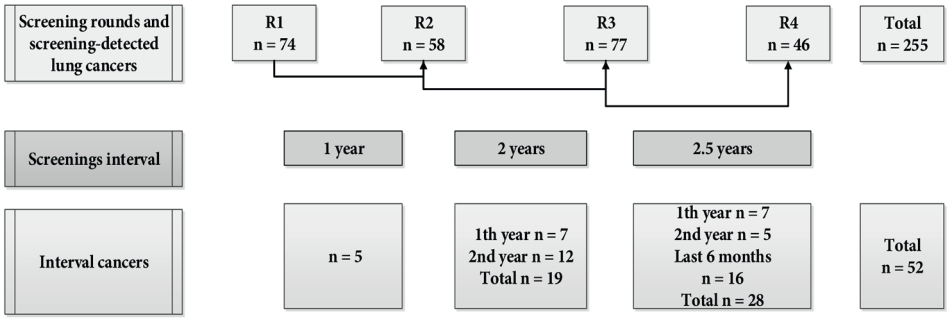


Figure 3: overview of the screening-detected lung cancers and interval cancers across the four rounds. The numbers of lung cancers presented are not equal to the number of participants with lung cancer: as 12 participants with screening-detected lung cancer (round 1 n = 4; round 2 n = 3; round 3 n = 2; and round 4 n =3), and one participant with an interval lung cancer (second year round 3) were diagnosed with synchronous double tumors.

2.4. DISCUSSION

In this study, the NELSON screening strategy of the final screening round was evaluated. Compared with the first three rounds a higher proportion of new lung cancers was detected at an advanced disease stage (stage IIIb/IV) and the interval cancer rate was higher in the 2.5-year interval compared with the 1-year and 2-year screening intervals.

Relative to the first three rounds, the participation rate in round 4 was slightly lower (80.7% vs 87.5–95.5%) [12]. One explanation could be that the original NELSON study protocol consisted of three screening rounds and in order to perform a fourth screening round an additional

informed consent was necessary. At that time, 6 years after randomisation, some participants were no longer traceable. Another explanation could be that at the time of dispatching the invitations for the fourth screening round, participants may have lost interest in further screening.

Participants with exclusively negative test results were more willing to participate in the additional screening round than participants with at least one non-negative screening result.

Moreover, eligible participants who were screened in the fourth round, were more often current smokers. This could support results of previous analyses in the NELSON trial, indicating that screening may have an unintended health certificate effect that permits continued smoking [19]. This indicates that lung cancer screening should be coupled with a smoking cessation intervention.

In the final round, 2.0% (105/5,380) of the screenings had a positive screening result. This resulted in a total of 2.0% (598/ 29,737) positive screenings in the NELSON study across all four rounds, which is comparable to the Danish Lung Cancer Screening Trial (DLCST; 2.0%) [20]. Compared with the NLST, the proportion of positive screenings in the NELSON trial is substantially lower (2.0% vs 24.2%) [3, 21]. At the same time, the NLST also reported a substantially higher FP rate after a positive screening than the NELSON trial (96.4% vs 59.4%) [3, 21].

In round 4, a lower proportion of screenings yielded an indeterminate or positive result than in round 3, which took place 2 years after the second round [12]. This could be due to the NELSON nodule management protocol allowing the radiologist to categorise abnormalities that remained stable across the previous rounds as negative [15, 17, 18]. Another influencing factor could be the finding that substantially more participants were diagnosed with an interval cancer in the 2.5-year interval compared with the 1-year and 2-year intervals, leading to fewer participants with suspicious abnormalities at the time of screening in round 4.

The cumulative lung cancer detection rate across the four rounds is 3.2%, which is comparable with the DLCST [22]. Relative to the NLST, the cumulative lung cancer detection rate of the NELSON trial is substantially higher: 3.2% vs 2.4%. However, the NLST had three annual screenings, a different nodule management protocol, and a different study population [3, 21]. The effectiveness of the NELSON trial (including the proportion of screening-detected lung cancers that are overdiagnosed) is yet to be determined.

Analysis of the first two intervals showed that a 2-year interval between the second and third screening rounds did not lead to significantly more advanced stage lung cancers compared with a 1-year interval between the first and second rounds ($p=0.09$) [11]. However, the fourth round led to a stage shift in screening-detected cancers that was significantly less favourable than after a 1-year screening interval (eg, more stage IIIb/IV cancers) [3, 20]. It also led to significantly higher proportions of SQM, BAC and SCLC ($p<0.001$). A higher proportion of

SQM and SCLC could be a result of more current smokers and a higher age at the moment of screening in round 4. However, the absolute numbers of these detected cancers were small. Compared with a 2-year screening interval, there was a similar tendency towards unfavourable change in stage distribution for a 2.5-year screening interval although this did not reach statistical significance. Also, the interval cancer rate was 1.47 (28/19) times higher in the 2.5-year interval compared with the 2-year interval. Moreover, in the last 6 months before the final fourth screening round the interval rate was 1.3 (16/12) times higher than in the first 24 months after the third round, suggesting that a 2.5-year interval may be too long.

On average, 69.4% of the screening-detected lung cancers across the four screening rounds in the NELSON trial were diagnosed in stage I and 9.8% in stage IIIb/IV [11]. This cumulative stage distribution of the screening-detected lung cancers in the NELSON trial appears to be favourable compared with those of the DLCST and the NLST (68.1% and 61.6% of cancers at stage I, and 15.9% and 20.0% at stage IIIb/IV, respectively) [3, 20]. However, this finding should be interpreted with caution because (1) the NLST used the 6th edition of the TNM (tumour, node, metastases) staging system, while the NELSON trial used the 7th edition, (2) the NLST and DLCST applied different eligibility criteria than the NELSON trial and (3) the proportion of overdiagnosed lung cancers in the screening group is yet unknown. The lung cancers found in the NELSON control group have yet to be investigated.

The strengths of this study include its population-based randomized setting, with a large number of participants in the screening and control groups. Second, by incorporating an indeterminate test outcome in the nodule management protocol instead of only two possible outcomes (eg, negative or positive), it seems possible to arrive at a better distinction between participants who might and who might not benefit from a diagnostic workup, leading to fewer FP results (ie, a better harm-benefit ratio). The limitations of this actual substudy were the relatively small absolute number of screening-detected lung cancers in the fourth round and the small absolute numbers of interval cancers between the third and fourth rounds. Furthermore, data on interval cancers after the fourth round were not yet available, and therefore no analyses of screening sensitivity of the final round could be performed.

In conclusion, a 2.5-year screening interval after the third round likely reduces the effectiveness of screening: in the final round significantly more advanced disease stage lung cancers were detected compared with a 1-year screening interval and compared with a 2-year screening interval a similar unfavourable change in stage distribution was seen, however not statistically significant. The proportion of interval cancers in the 2.5-year interval was substantially higher compared with a 1-year and a 2-year screening intervals. Modelling will give more insight into the potential effect of the different screening intervals in the NELSON trial.

ACKNOWLEDGEMENTS

The authors thank the system controllers R Faber and FJP Santegoets, and the secretary M Quak (all from the department of Public Health, Erasmus University Medical Center) for their contribution and maintenance of the database. Furthermore, the authors thank R Ziengs (University Medical Center Groningen) and S van Amelsvoort-van der Vorst (University Medical Center Utrecht). Finally, the authors thank the Dutch Cancer Registry (NKR) and the Vital Statistics (CBS) of the Netherlands for the data linkages.

SUPPLEMENTARY MATERIALS

Table 1: localisations and proportion female participants of screen-detected lung cancers in round 2, round 3 and round 4

	Round 2		p-value[1]	Round 3		p-value[2]	Round 4	
	n	%	0.91	n	%	0.66	n	%
Localisation								
Right lobes	37	63.8		48	62.3		31	67.4
Left lobes	21	34.5		29	37.7		15	32.6
Total	58	100		77	100		46	100
Females	10	17.2		10	13.0		7	15.2

[1]p-value: comparison between localization of screen-detected lung cancers of round 2 vs. round 4
[2]p-value: comparison between localization of screen-detected lung cancers of round 3 vs. round 4

Table 2: stage distributions of interval cancers detected in the 2.5 years interval between the third and fourth screening round

Interval cancers	Detected in the first 24 months	Detected in the last six months	p-value
Stage	n (%)	n (%)	0.77
Ia	2 (16.7)	1 (6.2)	
Ib	-	1 (6.2)	
IIa	-	1 (6.2)	
IIb	2 (16.7)	1 (6.2)	
IIIa	-	2 (12.5)	
IIIb	1 98.3)	3 (18.8)	
IV	7 (58.3)	7 (43.8)	
Total	12 (100)	16 (100)	

Table 3: histology of interval cancers detected in the 2.5 years interval between the third and fourth screening round

Interval cancers	Detected in the first 24 months	Detected in the last six months	p-value
	n (%)	n (%)	0.32
Histology			
Adenocarcinoma	4 (33.3)	5 (31.2)	
Squamous-cell carcinoma	1 (8.3)	4 (25.0)	
Large cell carcinoma	2 (16.7)	1 (6.2)	
SCLC	-	3 (18.8)	
Others	2 (16.7)	-	
No diagnosis possible	1 (8.3)	2 (12.5)	
NSCLC, not specified	1 (8.3)	1 (6.2)	
Large cell neuro-endocrine tumor	1 (8.3)	-	
Total	12 (100)	16 (100)	

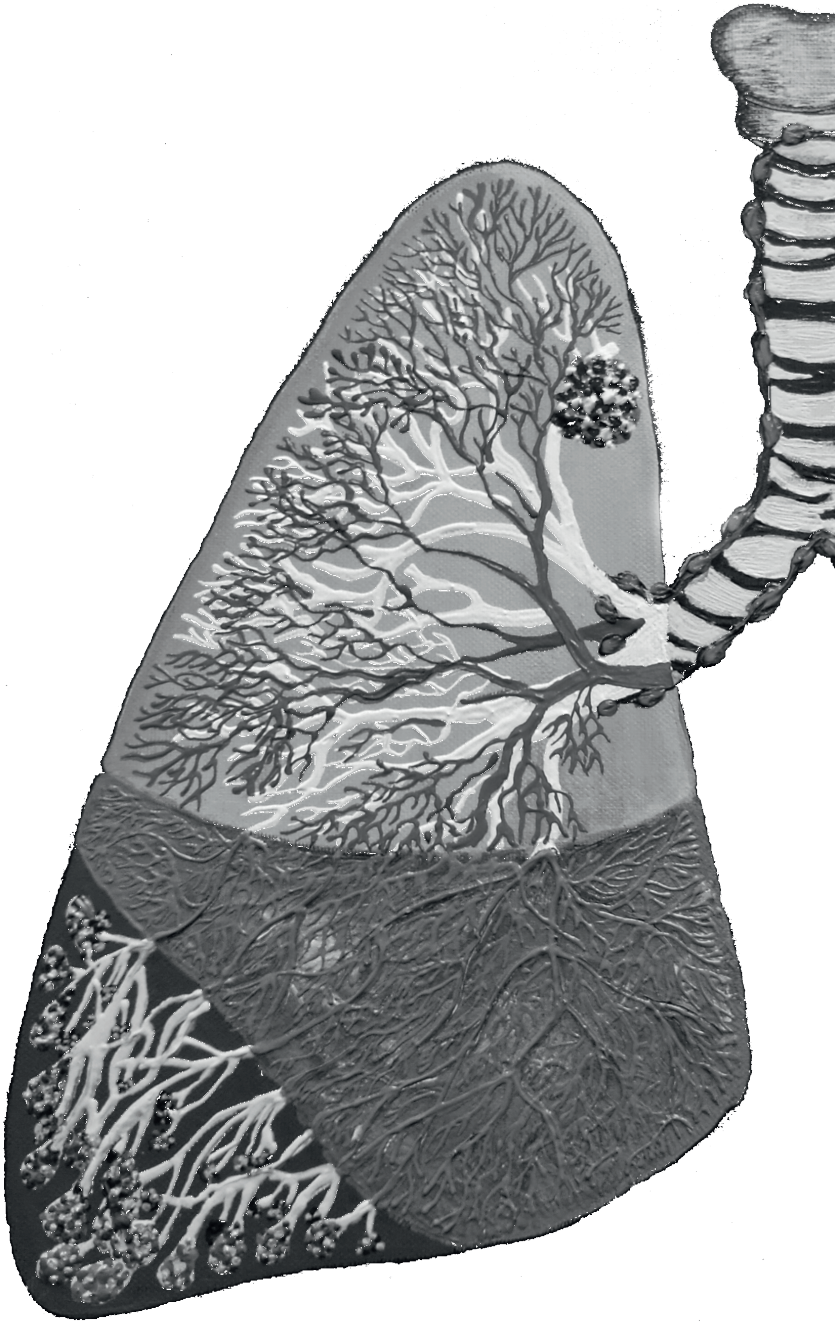
Table 4: proportion of female participants and current smokers in the participants with interval cancers detected in the 2.5 years interval between the third and fourth screening round

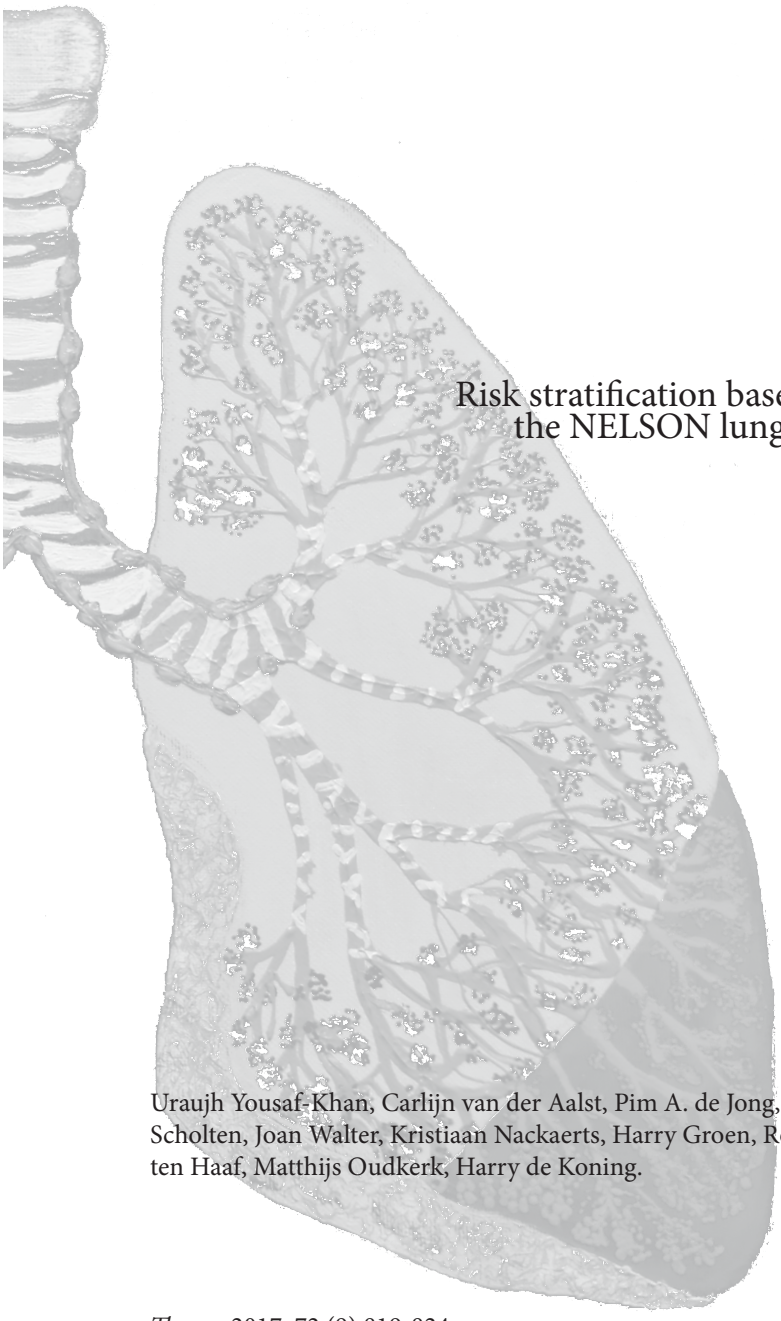
	First 24 months	Last 6 months	p-value
Female, n (%)	3 (25.0%)	1 (6.2%)	0.29
Current smokers, n (%)	10 (83.3%)	2 (16.7%)	0.22

REFERENCES

1. Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. CA Cancer J Clin. 2014;64(1):9-29.
2. Aberle Dr Fau - Berg CD, Berg Cd Fau - Black WC, Black Wc Fau - Church TR, Church Tr Fau - Fagerstrom RM, Fagerstrom Rm Fau - Galen B, Galen B Fau - Gareen IF, et al. The National Lung Screening Trial: overview and study design. (1527-1315 (Electronic)).
3. National Lung Screening Trial Research T, 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.
4. Moyer VA, Force USPST. Screening for lung cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2014;160(5):330-8.
5. 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.
6. 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. Eur J Cancer Prev. 2012;21(3):308-15.
7. Sverzellati N, Silva M, Calareso G, Galeone C, Marchiano A, Sestini S, et al. Low-dose computed tomography for lung cancer screening: comparison of performance between annual and biennial screen. Eur Radiol. 2016;26(11):3821-9.
8. van Iersel CA, de Koning HJ, Draisma G, Mali WP, 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.
9. 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 Med. 2009;361(23):2221-9.
10. Xu DM, Gietema H, de Koning H, Vernhout R, Nackaerts K, Prokop M, et al. Nodule management protocol of the NELSON randomised lung cancer screening trial. Lung Cancer. 2006;54(2):177-84.
11. Horeweg N, van der Aalst CM, Thunnissen E, Nackaerts K, Weenink C, Groen HJ, 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.
12. Horeweg N, van der Aalst CM, Vliegenthart R, Zhao Y, Xie X, Scholten ET, et al. Volumetric computed tomography screening for lung cancer: three rounds of the NELSON trial. Eur Respir J. 2013;42(6):1659-67.
13. Horeweg N, Nackaerts K, Oudkerk M, de Koning HJ. Low-dose computed tomography screening for lung cancer: results of the first screening round. J Comp Eff Res. 2013;2(5):433-6.
14. Horeweg N, Scholten ET, de Jong PA, van der Aalst CM, Weenink C, Lammers JW, et al. Detection of lung cancer through low-dose CT screening (NELSON): a prespecified analysis of screening test performance and interval cancers. Lancet Oncol. 2014;15(12):1342-50.
15. Xie X, Willemink MJ, Zhao Y, de Jong PA, van Ooijen PM, Oudkerk M, et al. Inter- and intrascanner variability of pulmonary nodule volumetry on low-dose 64-row CT: an anthropomorphic phantom study. Br J Radiol. 2013;86(1029):20130160.

16. Xie X, Zhao Y Fau - Snijder RA, Snijder Ra Fau - van Ooijen PMA, van Ooijen Pm Fau - de Jong PA, de Jong Pa Fau - Oudkerk M, Oudkerk M Fau - de Bock GH, et al. Sensitivity and accuracy of volumetry of pulmonary nodules on low-dose 16- and 64-row multi-detector CT: an anthropomorphic phantom study. (1432-1084 (Electronic)).
17. Wang Y, van Klaveren RJ, van der Zaag-Loonen HJ, de Bock GH, Gietema HA, Xu DM, et al. Effect of nodule characteristics on variability of semiautomated volume measurements in pulmonary nodules detected in a lung cancer screening program. Radiology. 2008;248(2):625-31.
18. Ru Zhao Y, 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.
19. van der Aalst CM, van Klaveren RJ, de Koning HJ. Does participation to screening unintentionally influence lifestyle behaviour and thus lifestyle-related morbidity? Best Pract Res Clin Gastroenterol. 2010;24(4):465-78.
20. 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.
21. Aberle DR, DeMello S Fau - Berg CD, Berg Cd Fau - Black WC, Black Wc Fau - Brewer B, Brewer B Fau - Church TR, Church Tr Fau - Clingan KL, et al. Results of the two incidence screenings in the National Lung Screening Trial. (1533-4406 (Electronic)).
22. Wille MM, Dirksen A, Ashraf H, Saghir Z, Bach KS, Brodersen J, et al. Results of the Randomized Danish Lung Cancer Screening Trial with Focus on High-risk Profiling. (1535-4970 (Electronic)).





CHAPTER III

Risk stratification based on screening history: the NELSON lung cancer screening study

Uraujh Yousaf-Khan, Carlijn van der Aalst, Pim A. de Jong, Marjolein Heuvelmans, Ernst Scholten, Joan Walter, Kristiaan Nackaerts, Harry Groen, Rozemarijn Vliegenthart, Kevin ten Haaf, Matthijs Oudkerk, Harry de Koning.

Thorax 2017; 72 (9):819-824

ABSTRACT

Background
Debate about the optimal lung cancer screening strategy is ongoing. In this study, previous screening history of the Dutch-Belgian Lung Cancer Screening trial (NELSON) is investigated on if it predicts the screening outcome (test result and lung cancer risk) of the final screening round.

Methods
15,792 participants were randomised (1:1) of which 7,900 randomised into a screening group. CT screening took place at baseline, and after 1, 2 and 2.5 years. Initially, three screening outcomes were possible: negative, indeterminate or positive scan result. Probability for screening outcome in the fourth round was calculated for subgroups of participants.

Results
Based on results of the first three rounds, three subgroups were identified: (1) those with exclusively negative results (n=3,856; 73.0%); (2) those with ≥1 indeterminate result, but never a positive result (n=1,342; 25.5%); and (3) with ≥1 positive result (n=81; 1.5%). Group 1 had the highest probability for having a negative scan result in round 4 (97.2% vs 94.8% and 90.1%, respectively, $p<0.001$), and the lowest risk for detecting lung cancer in round 4 (0.6% vs 1.6%, $p=0.001$). 'Smoked pack-years' and 'screening history' significantly predicted the fourth round test result. The third round results implied that the risk for detecting lung cancer (after an interval of 2.5 years) was 0.6% for those with negative results compared with 3.7% of those with indeterminate results.

Conclusions
Previous CT lung cancer screening results provides an opportunity for further risk stratifications of those who undergo lung cancer screening.

3.1. INTRODUCTION

Lung cancer is the leading cause of cancer-related death worldwide [1]. Lung cancer is often diagnosed at an advanced disease stage and occurs increasingly among former smokers [2]. This underlines the need for preventive measures. Since 2013, lung cancer screening has been recommended by the United States Preventive Services Task Force [3, 4]. People aged 55 through 80, who have smoked at least 30 pack-years, and currently smoke or have quit smoking within the past 15 years are invited for an annual low-dose CT (LDCT) examination in the USA. However, debate about the optimal screening strategy (eg, the optimal screening interval) is still ongoing [5-9].

Currently, lung cancer screening is not implemented in Europe. The main reason is that none of the (underpowered) European lung cancer screening trials have shown a mortality reduction so far [10-13]. However, Europe's largest trial the Dutch-Belgian lung cancer screening trial (NELSON) is now in its final follow-up phase and is sufficiently powered to detect a lung cancer mortality reduction of at least 25% [14-16]. Main differences between NELSON and the National Lung Screening Trial (NLST) are as follows: (1) the age of the selected subjects (50–74 vs 55–74 years), (2) the use of increasing screening intervals versus annual screening, (3) volumetric-based nodule management versus a diameter-based nodule management and (4) a control group with no screening versus a control group screened with annual chest radiography. The differences in screening interval enable to investigate the optimal screening strategy for lung cancer screening to reduce the probability on potential harms (eg, false-positive examinations leading to unnecessary (non)-invasive diagnostic procedures) for those without lung cancer [4, 17, 18] The NLST showed that lung cancer risk and mortality benefit vary within the screened population: the largest mortality benefit was achieved in the subgroup with the highest risk for developing lung cancer [4]. The ratio between benefits and harms of lung cancer screening could be improved by more precisely identifying a high-risk population for developing lung cancer [19-21] and by risk stratification of subjects based on the individual's screening history (eg, previous screening outcome or presence of a nodule) [6 22-24]. Recently, the NLST showed that participants with a negative screening result at baseline have a lower lung cancer detection risk at subsequent screening rounds compared with all screened participants and therefore may not need annual lung cancer screening [7]. In line with this, the baseline scan in the NELSON trial allowed the identification of three subgroups with different risks for detecting lung cancer in the second and third screening rounds [23]. Since all NELSON screening rounds with different screening intervals (with a unique 2.5-year interval) have been completed, this study aims to investigate whether NELSON subgroups with different risks for detecting lung cancer can be identified based on their previous screening history. This information might be useful for further risk stratification of subjects who undergo lung cancer screening.

3.2. METHODS

NELSON trial

In brief, the NELSON trial is a randomised-controlled, population-based lung cancer screening trial. The primary aim is to the investigate whether LDCT screening of high-risk subjects for developing lung cancer can lead to a reduction of lung cancer mortality by ≥25% compared with no screening at 10 years of follow-up [15]. High-risk subjects, mainly males, were defined as aged between 50 and 74 years, who had smoked at least 15 cigarettes/day for ≥25 years or 10 cigarettes/day for ≥30 years, and were still smoking or had quit <10 years ago [14].

Initially, 15,822 participants were randomised (1:1). However, through linkages with the national cancer registries and death registries of the Netherlands and Belgium (Statistics Netherlands and the Flemish Agency for Care and Health, respectively), it appeared that 30 participants (15 screening group and 15 control group participants) died before randomisation and should therefore be ruled out from further analysis. After this correction, 15,792 participants were randomised (1:1) into a CT screening group (7,900) and into a control group (n=7,892). Screening took place at baseline, after 1 year, 3 years and 5.5 years. The control group received usual care (no screening). The baseline screening round was conducted from January 2004 through December 2006, and the final screening round was conducted from November 2009 through March 2012.

Study participants

For the risk stratifications based on the regular scan results of the first three screening rounds, only participants who attended both one of the first three regular screening rounds and round 4 were included (n=5,279). For the risk stratification based on the results of the third screening round alone, participants who attended both the third and fourth rounds were included (n=5,268). To compare participants with and without a screening-detected lung cancer across all the screening rounds, all screened participants were included (n=7,582).

Screening procedures, outcomes and the nodule management protocol

Screening was performed using 16-detector CT scanners in low-dose setting at four screening sites (University Medical Center Groningen, University Medical Center Utrecht, Kennemer Gasthuis Haarlem in the Netherlands and University Hospital Gasthuisberg Leuven in Belgium) [25]. More detailed descriptions of the equipment, the execution of the screening examination and the nodule management protocol have been published previously [15, 23, 25-27]. In short, screening could lead to the following test results: (1) negative: no nodule, newly detected nodule with a volume of <50 mm3 or previously detected nodule with a growth (change in volume between scans) of <25% or ≥25% but with a volume doubling time (VDT) of >600 days; (2) indeterminate: newly detected nodule with a volume of 50–500 mm3 or

previously detected nodule with a VDT of 400–600 days; or (3) positive: newly detected nodule with a volume of >500 mm3 or previously detected nodule with a VDT of <400 days [15].

Participants with a negative screening result were invited for the next regular screening round. Those with an indeterminate screening result underwent a low-dose follow-up CT scan to measure volume growth and VDT after 6 weeks to 4 months or after 12 months (depending on nodule volume and screening rounds), to define their definitive screening result (negative or positive). Those with a positive screening result were referred to a pulmonologist for a diagnostic work-up. If no lung cancer was diagnosed, the participant was referred to the next regular screening round. If lung cancer was confirmed, the patient received treatment according to the Dutch National guidelines. Medical data of these patients were collected prospectively.

Definitions

Current smokers were those who were smoking or who had smoked in the last 7 days before completion of the baseline (risk) questionnaires. Former smokers have quit smoking for 10 years or less. Variables that were calculated were pack-years (20 cigarettes smoked per day for 1 year) and body mass index (BMI) (body weight (in kilograms)/the square of body length (in meters)). New nodules were nodules which were labelled as 'new nodule' or labelled as 'not new, but too small in previous scan to be detectable' by radiologists [24].

Regular round scans were the first CT examinations in a regular screening round (years 1, 2, 4 and 6.5). Follow-up scans were repeat LDCT scans after an Indeterminate result in one of the four regular rounds. A screening-detected lung cancer was defined as a lung cancer diagnosed by a pulmonologist within 12–24 months, depending on the screening round, after referral for a positive screening. Lung cancer detection rate was defined as the number of screening-detected lung cancers divided by the number of screened participants. Regular scan result was defined as the result of the first CT examination in a screening round, while the definitive outcome of the screening (screening result) was made after inclusion of the results of the follow-up scans at the conclusion of that screening round.

Risk groups

Three unique subgroups were identified based on regular scan results of the first three screening rounds: (1) participants with solely negative results (n=3,856; 73%); (2) participants with ≥1 indeterminate result and never a positive result (n=1,342, 25.5%); and (3) participants with ≥1 positive result (n=81, 1.5%; online supplementary **table S1**).

Based on the regular scan results of round 3, three other unique subgroups were identified: (1) those with a negative result in round 3 (n=4,925, 93.5%); (2) those with an indeterminate result in round 3 (n=324, 6.2%); and (3) those with a positive result in round 3 (n=19, 0.4%).

Statistical analyses

None of the continuous variables were distributed normally, tested by using the Kolmogorov-Smirnov test and examining Q–Q plots. Therefore, the variables were described by using medians and IQRs. Differences between the continuous variables across the subgroups were calculated by using the median test or analysis of variance, depending on number of subgroups. Differences between nominal variables were calculated by using a $\chi 2$ test. Differences between categorical variables were tested by using a Mann-Whitney U test or Kruskal-Wallis test, depending on the number of subgroups. To test differences between the subgroups regarding the probability of having a negative or nonnegative (indeterminate or positive test result) in the fourth round, and to test differences between the subgroups regarding the risk for detecting lung cancer in the fourth round, a logistics regression analysis was used. The associations of gender, pack-years smoked, age at randomisation, Smoking status at randomization and age of starting smoking with the screening outcome of the fourth round were assessed through univariate and multivariate analyses. For multivariate analyses, an ordinal logistic regression model was developed using backward selection. In addition, variables were also tested for interactions. For all analyses, a p value <0.05 was considered significant, and IBM SPSS Statistics V.21 was used.

3.3. RESULTS

In the first three rounds, 7,557, 7,295 and 6,922 participants were screened, respectively [23]. In round 4, 5,279 participants were screened and 5,380 scans were performed; of which 5,279 were regular scans and 101 follow-up scans [16]. In this section, the probabilities of screening outcomes and risk for detecting lung cancer in round 4 were calculated based on the regular scan results alone. An overview of these probabilities based on the definitive screening results of round 4 is presented in the online data supplement (see **Table S1**).

Risk stratification based on the results of the first three screening rounds

Based on the regular scan results of the first three rounds, three subgroups were identified: (1) participants with only negative results; (2) participants with ≥1 indeterminate, but never a positive result; and (3) participants with ≥1 positive result (**Table 1**). Those from group 1 were slightly younger (57.0 vs 58.0 years, p<0.001) compared with the participants of the other subgroups. Participants with ≥1 positive result had smoked statistically significant slightly more pack-years than the other subgroups (p=0.02). No significant differences were observed between gender and baseline smoking status among the subgroups (p=0.66 and p=0.23, respectively).

Participants with ≥1 indeterminate result, but never a positive result (OR 1.89, p=0.001), and participants with ≥1 positive result (3.77, p<0.001; **Table 2**) had a significantly higher OR of receiving a non-negative scan result (e.g., an indeterminate or a positive result) in round 4

compared with the group with solely negative results in the first three rounds (the reference group).

In univariate analysis, only the screening history (p<0.001) and smoked pack-years (p=0.01) significantly predicted the regular scan result in round 4, while gender (p=0.70), age (p=0.13) and baseline smoking status (p=0.75) did not. In multivariate analysis, screening history and smoked pack-years remained significant predictors (p<0.001 and p=0.02, respectively): the model suggests that screening history and pack-years are positively associated with the screening outcome in the fourth round (see online supplementary **Table S4**). Interaction between subgroups and smoked pack-years was not significant (p=0.89).

In round 4, 43 participants were diagnosed with 46 screening-detected lung cancers. OR for detecting lung cancer in round 4 differed between the subgroups as well: relative to the group with only negative results, the group with ≥1 indeterminate result (but never a positive result) had an OR of 2.77 (p<0.001) for detecting lung cancer in round 4 (**Table 3**). No lung cancer was detected in the group with ≥1 positive result. None of the following factors predicted significantly the detection of lung cancer in round 4: gender (p=0.71), age (p=0.10), starting age of smoking (p=0.20), smoking status (0.28) or pack-years smoked (0.09; all data not shown). Multivariate analysis showed no statistically significance for age and pack-years smoked.

A total of 22 (51.2%) of the participants with screen-detected lung cancer in round 4 had solely negative scan results in the first three screening rounds, and in 20 (90.9%) of those participants lung cancer was detected in a new nodule (data not shown). The remaining 21 (48.8%) screen-detected lung cancers in round 4 were detected among participants with at least one indeterminate scan result in the previous three screening rounds, and in 12 (57.1%) of those participants lung cancer was detected in a new nodule (data not shown).

Participants with only negative regular scan results were stratified by pack-years smoked. Therefore, six categories were made: <25, 26–30, 31–35, 36–40, 41–45 and >45 years. For each category, the risk for lung cancer detection in the fourth round was calculated. The first five categories had a lung cancer risk between 0.2% (the first category) and 0.7% (the fifth category). In other words, the risk for detecting lung cancer in the fourth round for those who smoked less than 45 pack-years was lower than 0.7%. For those who smoked more than 45 pack-years (n=1091), lung cancer detection rate in round 4 was 1.1% (p=0.04). No correlation was observed between pack-years smoked and lung cancer detection rate in the final round (all data not shown) for the other subgroups.

Table 1: baseline characteristics of subgroups based on the previous screening rounds and of all participants of the fourth screening round (based on regular scan results)

Baseline characteristics	**Subgroups based on the results of the previous three screening rounds***			**All participants of the fourth screening round**	
	All negatives	≥1x indeterminate, but never a positive result	≥1x positive		p-value†
Male, n/N (%)	3,237/3,856 (83.9)	1,129/1,342 (84.1)	71/81 (87.7)	4,437/5,279 (84.1)	0.66
Age, median (IQR)	57.0 (7.0)	58.0 (7.0)	58.0 (7.0)	58.0 (8.0)	<0.001
Current baseline smokers, n/N (%)	2,086/3,856 (54.1)	753/1,342 (56.1)	39/81 (48.1)	2,878/5,279 (54.5)	0.23
Pack-years at baseline, median (IQR)	38.0 (19.8)	38.0 (19.8)	38.7 (23.5)	38.0 (19.8)	0.02
Total, n (%)	3,856 (73.0)	1342 (25.5)	81 (1.5)	5,279 (100.0)	-

*only participants screened in across at least the first three rounds as well as in the fourth screening round were included

† p-value across the subgroups by baseline characteristics

Table 2: risk calculation for subgroups based on the regular scan results of the first three rounds

	Regular round scan results of the fourth screening round										
	Total group		Negative		Indeterminate		Positive		Non-negative result*		p-value
	N	%	n	%	n	%	n	%	OR	95% CI	
Subgroups											
All negatives	3,856	73.0	3,747	97.2	61	1.6	48	1.2	REF†	REF†	NA
≥1x indeterminate, but never a positive result	1,342	25.5	1,272	94.8	37	2.8	33	2.5	1.9	1.4-2.6	0.001
≥1x positive result	81	1.5	73	90.1	3	3.7	5	6.2	3.8	1.8-8.0	<0.001
All participants screened in round 4	5,279	100.0	5,092	96.5	101	1.9	86	1.6	-	-	-

*Non-negative result means an indeterminate or a positive scan result (based on logistic regression)

†The subgroup with only negative scan results in the first three rounds was the reference group to calculate the odds ratio.

Risk stratifications based on the previous screening round

Round 4 was performed 2.5 years after the third screening round and 5.5 years after the baseline scan. Participants with a negative scan result in round 3 were significantly younger at baseline than those with an indeterminate or a positive third round result (57.0 years vs 59.0 and 59.0 years, p<0.001; **Table 4**). The probability for detecting lung cancer in the fourth round differed between participants with a negative scan result and an indeterminate scan result: 0.6% vs 3.7%, respectively (p<0.001; **Tables 5 and 6**). No lung cancer was detected in round 4 in the small group with a positive scan result in round 3.

Participants with a screen-detected lung cancer

Across all four screening rounds, 243 out of 7,582 participants were diagnosed with a total of 255 screening-detected lung cancers. Participants with a screening-detected lung cancer were significantly older (61 vs 58 years, p<0.001) and had smoked more pack-years (44.0 vs 38.0 years, p<0.001) than those without a screening-detected lung cancer. Of those with screen-detected lung cancers, 28.4% had ≥1 indeterminate scan result (but never a positive test result initially) and 71.6% had ≥1 positive scan result (this group also contains those with once a negative or an indeterminate result) before diagnoses of lung cancer. No differences were seen in gender (p=0.98), baseline smoking status (p=0.61) or in BMI between participants with or without screening-detected lung cancer (p=0.38; all data not shown).

Table 3: risk to detect lung cancer in round 4 for subgroups based on the regular scan results of the first three rounds

	Screening-detected lung cancer in round 4						
	No screening-detected lung cancer		Yes, a screening- detected lung cancer				p-value
	n	%	N	%	OR*	95% CI	
Subgroups							
All negatives	3,834/3,856	99.4	22/3,856	0.6	REF†	REF†	NA
≥1x indeterminate, but never a positive result	1,321/1,342	98.4	21/1,342	1.6	2.77	1.52-5.05	0.001
≥1x positive result	81/81	100.0	-	-	NA	NA	NA
All participants screened in round 4	5,236/5,279	99.2	43/5,279	0.8	-	-	-

*The risk for detecting lung cancer in the fourth screening round expressed in odds ratio (based on logistic regression)

†The subgroup with only negative scan results in the first three rounds was the reference group to calculate the odds ratio.

Table 4: baseline characteristics of subgroups based on the regular scan results of the third screening round alone

Baseline characteristics	Subgroups based on the screening results of the third round*			
	Negative	Indeterminate	Positive	p-value†
Male, n/N (%)	4,143/4,925 (84.1)	270/324 (83.3)	16/19 (84.2)	0.93
Age, median (IQR)	57.0 (8.0)	59.0 (8.0)	59.0 (7.0)	<0.001
Current smokers, n/N (%)	2,672/4,925 (54.3)	195/324 (60.2)	7/19 (36.8)	0.04
Pack-years, median (IQR)	38.0 (19.8)	38.7 (19.8)	43.2 (30.0)	0.05

*only participants screened in both the third and fourth round were included.
†p Value across the subgroups by baseline characteristics

Table 5: risk calculation for subgroups based on the regular scan results of round three alone

Result round three	**Regular round scan results of the fourth screening round**										
	Total group		Negative		Indeterminate		Positive		non-negative result*		p-value
	n	%	n	%	n	%	n	%	OR	95% CI	
Negative	4,925	93.5	4,770	96.9	88	1.8	67	1.4	REF†	REF†	NA
Indeterminate	324	6.2	293	90.4	12	3.7	19	5.9	3.3	2.2-4.9	<0.001
Positive	19	0.4	18	94.7	1	5.3	-	-	1.7	0.2-12.5	0.60
Total	5,268	100.0	5081	96.5	101	1.9	86	1.6	-	-	-

*Non-negative result means an indeterminate or a positive scan result.
†The group with a negative scan result in round three was the reference group to calculate the odds ratio (based on logistic regression).

Table 6: risk to detect lung cancer in round four for subgroups based on the regular scan results of round three alone

Subgroups	**Screen-detected lung cancer in round four**						
	No screening-detected lung cancer		Yes, a screening- detected lung cancer				p-value
	n	%	n	%	OR*	95% CI	
Negative	4,894/4,925	99.4	31/4,925	0.6	REF†	REF†	
Indeterminate	312/324	96.3	12/324	3.7	6.1	3.1-11.9	<0.001
Positive	19/19	100	-	-	NA	NA	NA

*The risk for detecting lung cancer in the fourth screening round expressed in odds ratio.
†The group with negative scan result in round three was the reference group to calculate the odds ratio (based on logistic regression).

3.4. DISCUSSION

This study demonstrated that individual's screenings history can be used as a risk stratification tool for their further screening regime. The probability for screening outcome in the fourth round differs across previous screening test result(s). Also, the risk for detecting lung cancer in the fourth screening round differs based on the previous screening outcome(s).

Previous NELSON results showed that the risk for detecting lung cancer in the subsequent second and third rounds differed among the baseline scan result [23]. Those with an indeterminate or a positive baseline scan result had a higher risk for detecting lung cancer in round 2 or 3, compared with those with a negative baseline scan result. In this study, the results indicated that the probability for non-negative (ie, indeterminate or positive) scan result in the fourth round was higher for those with ≥1 indeterminate (but never a positive result) and those with ≥1 positive result in the first three rounds, compared with those with only negative results in the first three rounds. Furthermore, it was demonstrated that having an older age and have smoked more pack-years smoked were both significant predictors for non-negative result. These results were in line with our previous study results [23]. None of the lung cancers detected in round 4 was detected in the group with previously ≥1 positive scan result. Moreover, the risk for detecting lung cancer in the final fourth round was non-significantly higher for those with an indeterminate or a positive definitive screening result compared with solely negative screening results (OR of 2.95 and 2.10, respectively).

The combined results of the previous screening rounds turned out to predict the screening outcome (scan results or lung cancer detection risk) in the fourth screening round. The third round test result predicted the fourth round test result after an interval of 2.5 years; for those with a negative third scan, a subsequent round with a 2.5 years interval seemed even short, as the lung cancer risk was <1% (as across all screening rounds) [16] [23]. Moreover, those with previous solely negative scan results and those with a third negative scan result may not need to be screened for more than 2.5 years, as the lung cancer detection rates were <0.7% or 1.1% at most for the fourth round, respectively. However, in almost 90%of those with solely negative scan results, the lung cancer was detected in a new nodule. Although malignant new nodules might be fast growing, detection at early stage with LDCT seems possible [24].

Furthermore, it was showed that having an indeterminate scan result gives a higher risk for a non-negative scan result and a higher risk for lung cancer detection in the final round. However, only a minority of the indeterminate nodules turn out to be malignant. With the growing evidence, cut-off points of the nodule management should be evaluated regularly to further optimise the ratio between benefits (eg, mortality reduction) and harms (e.g., false positive, unnecessary work-up) of the protocol. Moreover, cancer can evolve from nodules which are not seen before on the scan (e.g., from new nodules). One explanation could be 'field cancerisation', in which it is assumed that large areas of the bronchial epithelium are affected by smoking, leading to areas with metaplasia and dysplasia which sometimes turn out to be cancer and sometimes not. Moreover, it is known that in heavy smokers nodules appear

and disappear and come up in different pulmonary areas and are not always malignant.

Although only the NELSON trial used increasing screening intervals, the NLST showed in a recent retrospective analysis that participants with a negative baseline result had a lower incidence of lung cancer at baseline as well as a lower lung cancer detection rate in the subsequent rounds (0.34%) compared with all screened participants (1.0%).7 Furthermore, the lung cancer incidence and mortality for those with solely negative screening results was even lower than for those with a negative baseline screening. Their findings suggested that for the larger part of the screened population, it may lead to a better harm–benefit ratio to offer risk-based incidental screening rounds to participants with different screening intervals. In the NELSON trial, end results and cost-effectiveness analyses, and therefore the harm–benefit ratios of screening scenarios, are yet unknown. However, the current study concludes that previous screening history seems to be useful for risk stratification and to refine the screening protocol for subgroups with different risks for lung cancer [23, 24, 28].

Major strengths of this study are the large-scale, population-based randomised study design and its volumetric-based nodule management, leading to three initial screening outcomes. However, in this substudy, small numbers of screening-detected lung cancers were found in the fourth round and some subgroups had small numbers of participants. Furthermore, a subselection of screened participants was used: participants should have been screened in one of the three screening rounds as well as in the final screening round. Additionally, almost 1500 participants were lost to follow-up (no actual addresses) in the fourth round, since additional informed consent was required to perform the screening round [16].

In conclusion, the screening test result(s) might have major implications on the total number of scans needed for those who undergo lung cancer screening. This is useful for the further optimisation of the harm–benefit ratio of a lung cancer screening programme.

ACKNOWLEDGEMENTS

The authors thank the system controllers R. Faber and F.J.P. Santegoets, and the secretary M.Quak (all from the department of Public Health, Erasmus University Medical Center Rotterdam) for their contribution and maintenance of the database. Furthermore, the authors thank R. Ziengs (University Medical Center Groningen) and S. van Amelsvoort-van der Vorst (University Medical Center Utrecht). Finally, the authors thank the Dutch Cancer Registry (NKR) for the data linkages.

SUPPLEMENTARY MATERIALS

Table 1: possible combinations of regular scan results across the first three rounds

Possible combinations	Round 1	Round 2	Round 3	N	Subgroup
	N	N	N	3,826	ALL NEGATIVE
	X	N	N	15	ALL NEGATIVE
	N	X	N	6	ALL NEGATIVE
	N	N	X	9	ALL NEGATIVE
Total per subgroup					3,856
	I	I	I	17	≥1X INDETERMINATE, NEVER POSITIVE
	I	I	N	64	≥1X INDETERMINATE, NEVER POSITIVE
	I	N	I	78	≥1X INDETERMINATE, NEVER POSITIVE
	I	N	N	773	≥1X INDETERMINATE, NEVER POSITIVE
	N	I	N	180	≥1X INDETERMINATE, NEVER POSITIVE
	N	N	I	197	≥1X INDETERMINATE, NEVER POSITIVE
	N	I	I	22	≥1X INDETERMINATE, NEVER POSITIVE
	X	I	N	4	≥1X INDETERMINATE, NEVER POSITIVE
	X	I	X	1	≥1X INDETERMINATE, NEVER POSITIVE
	I	X	N	5	≥1X INDETERMINATE, NEVER POSITIVE
	I	N	x	1	≥1X INDETERMINATE, NEVER POSITIVE
Total per subgroup					1,342
	N	I	P	1	≥1X POSITIVE
	N	P	I	1	≥1X POSITIVE
	N	P	N	13	≥1X POSITIVE
	N	N	P	6	≥1X POSITIVE
	N	X	P	1	≥1X POSITIVE
	I	P	N	11	≥1X POSITIVE
	I	N	P	9	≥1X POSITIVE

Table 1: continued. Possible combinations of regular scan results across the first three rounds

Possible combinations	Round 1	Round 2	Round 3	N	Subgroup
	I	I	P	2	≥1X POSITIVE
	I	P	I	5	≥1X POSITIVE
	P	N	N	24	≥1X POSITIVE
	P	I	N	4	≥1X POSITIVE
	P	I	I	2	≥1X POSITIVE
	P	N	I	1	≥1X POSITIVE
	P	X	I	1	≥1X POSITIVE
Total per subgroup					81
TOTAL					5,279

X: In case a participant did not receive screening in a screening round
N = negative screening result
I = indeterminate screening result
P = positive screening result

Table 2: baseline characteristics of subgroups based on the previous definitive screening outcomes and of all participants of the fourth screening round

Baseline characteristics	Subgroups based on the results of the previous three screening rounds			All participants of the fourth screening round	
	All negatives	≥1x indeterminate, but never a positive result	≥1x positive[1]		p-value[2]
Male, n/N (%)	4,299/5,112 (84.1)	36/44(81.8)	102/123 (82.9)	4,437/5,279 (84.1)	0.86
Age, median (IQR)	57.0 (8.0)	58.4 (8.0)	60.0 (9.0)	58.0 (8.0)	<0.001
Current baseline smokers, n/N (%)	2,789/5,112 (54.6)	24/22 (54.5)	65/123 (52.8)	2,878/5,279 (54.5)	0.93
Pack-years at baseline, median (IQR)	38.0 (19.8)	38.0 (23.6)	38.7 (24.0)	38.0 (19.8)	0.06
Total, n (%)	5,112 (96.8)	44 (0.8)	123 (2.3)	5,279 (100.0)	-

[1]only participants screened across at least in of the first three rounds as well as in the fourth screening round were included

[2] p-value across the subgroups by baseline characteristics

Table 3a: risk calculation for subgroups based on the definitive screening results of the first three rounds

	Definitive screening results of the fourth screening round								
	Total group		Negative		Positive		Non-negative result[1]		p-value
	n	%	n	%	n	%	OR	95% CI	
Subgroups									
All negatives	5,105	100	5,015/5105	98.2	90/5,105	1.8	REF[2]	REF[2]	
≥1x indeterminate, but never a positive result	44	100	41/44	93.2	3/44	6.8	4.08	1.24-13.41	<0.001
≥1x positive result	123	100	111/123	90.2	12/123	9.8	6.03	3.21-11.32	0.02
All participants screened in round 4	5,272[2]	100.0	5,167/5,272	98.0	105/5,272	2.0			-

[1]Non-negative result means an indeterminate or a positive screening result

[2]Instead of 5,279 participants, in this analysis 7 participants were excluded as their definitive screening outcome was 'indeterminate', because they didn't show up for a follow-up scan.

Table 3b: risk to detect lung cancer in round four for subgroups based on the definitive screening results of the first three rounds

	Screening-detected lung cancer in round four						
	No screening-detected lung cancer		Yes, a screening- detected lung cancer				p-value
	n	%	N	%	OR[1]	95% CI	
Subgroups							
All negatives	5072/5112	99.2	40/5112	0.8	REF[2]	REF[2]	NA
≥1x indeterminate, but never a positive result	43/44	97.7	1/44	2.3	2.95	0.40-21.9	0.29
≥1x positive result	121/123	98.4	2/123	1.6	2.10	0.50-8.77	0.31
All participants screened in round 4	5,279	100	43[3]/5,279	100	-	-	-

[1]The risk for detecting lung cancer in the fourth screening round expressed in odds ratio.

[2]The group with a negative screening result in round three was the reference group to calculate the odds ratio.

[3]Of the 43 participants detected with screen detected lung cancer in round 4, at least 32 (74.4%) were detected with lung cancer in a new nodule (in 10 [23.3%] participants the cancer was detected in a nodule not registered as new or below screening limit on previous scans, and in one [2.3%] participant retrospective nodule matching proved impossible and no statement can be made). A total of 40 (93%) of the participants with screen detected lung cancer in round 4 had negative overall screening results in the first three rounds, and in 30 (75.0%) of those participants lung cancer was detected in a new nodule. In 27 (86.7%) of the 32 participants with new nodule lung cancer, the new nodule was observed in round 4 for the first time.

Table 4: final multivariate model resulting from backward stepwise selection

Parameter	Coefficient	Exp (coefficient)	95 CI	p-value
Screening history				
All negatives	REF	REF	REF	NA
≥1x indeterminate, but never a positive result	0.6	1.9	1.4-2.6	<0.001
≥1x positive result	1.3	3.8	1.8-8.0	0.001
Smoked pack-years	0.09	1.09	1.00-1.01	0.02

CI: confidence interval
REF: the group with all negatives was the reference group in the variable screening history
First, the univariate effect of the risk factors on the screening outcome in round four was assessed. Next, backward stepwise selection was used for variable selection in the multivariate model, using a 5% significance level as stopping criteria (based on Likelihood Ratio tests). Table 3 shows only the most important risk factors for screening outcome in round four.

REFERENCES

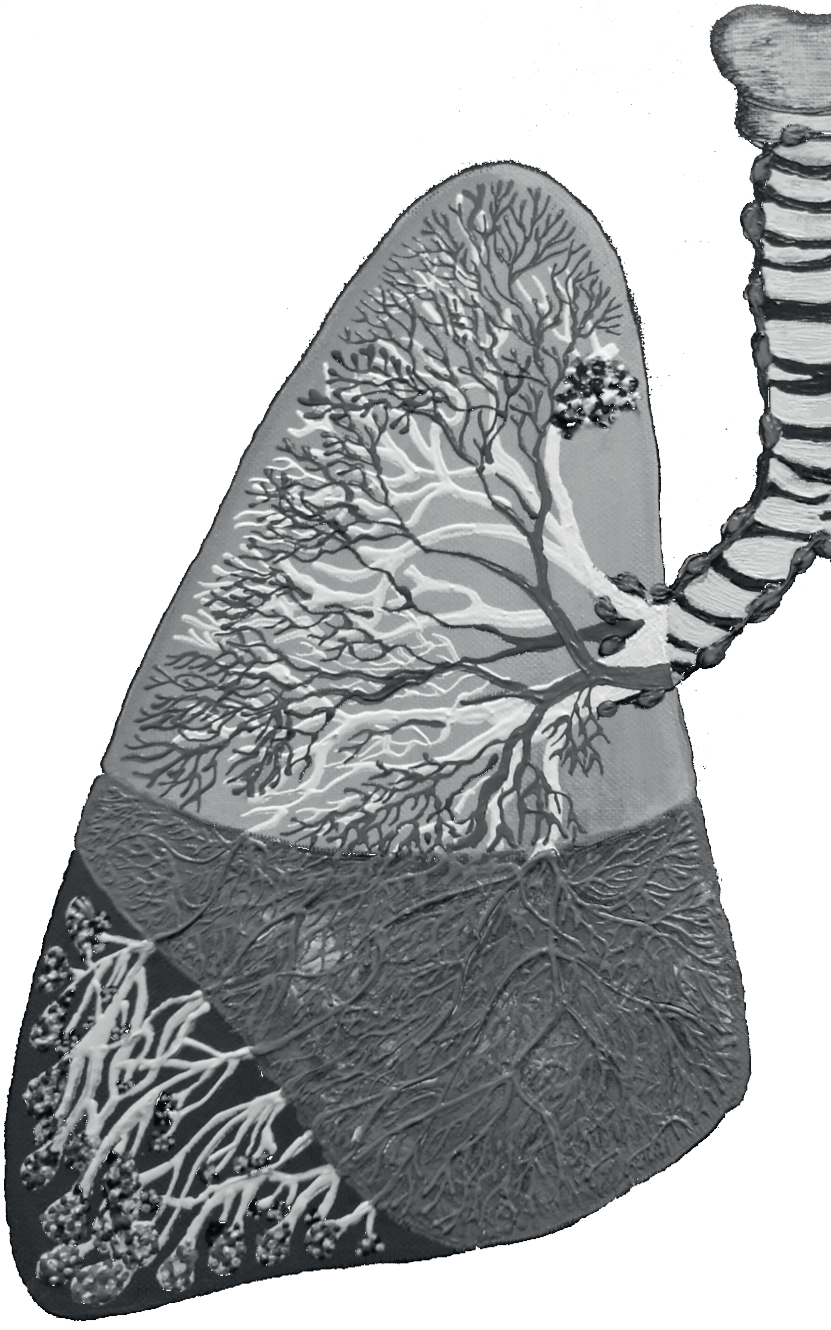
1. Torre LA, Siegel RL, Jemal A. Lung Cancer Statistics. Adv Exp Med Biol. 2016;893:1-19.
2. van Rens MT, Zanen P, Brutel de La Riviere A, *al. e*. Survival in synchronous vs. single lung cancer: upstaging better reflects prognosis. Chest. 2000;118(0012-3692 (Print)):952-8.
3. Moyer VA, Force USPST. Screening for lung cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2014;160(5):330-8.
4. National Lung Screening Trial Research T, 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.
5. Field JK, Duffy SW. Lung cancer CT screening: is annual screening necessary? Lancet Oncol. 2016;5(1474-5488 (Electronic)):543-4.
6. Horeweg N, Scholten ET, de Jong PA, van der Aalst CM, Weenink C, Lammers JW, et al. Detection of lung cancer through low-dose CT screening (NELSON): a prespecified analysis of screening test performance and interval cancers. Lancet Oncol. 2014;15(12):1342-50.
7. Patz EF, Jr., Greco E, Gatsonis C, Pinsky P, Kramer BS, Aberle DR. Lung cancer incidence and mortality in National Lung Screening Trial participants who underwent low-dose CT prevalence screening: a retrospective cohort analysis of a randomised, multicentre, diagnostic screening trial. Lancet Oncol.May;17(5)(1474-5488 (Electronic)):590-9.
8. Baldwin DR, Callister ME. What is the Optimum Screening Strategy for the Early Detection of Lung Cancer. (1433-2981 (Electronic)).
9. van der Aalst CM, Ten Haaf K, de Koning HJ. Lung cancer screening: latest developments and unanswered questions. (2213-2619 (Electronic)).
10. 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.
11. Wille MM, Dirksen A, Ashraf H, Saghir Z, Bach KS, Brodersen J, et al. Results of the Randomized Danish Lung Cancer Screening Trial with Focus on High-risk Profiling. (1535-4970 (Electronic)).
12. Infante M, Cavuto S Fau - Lutman FR, Lutman Fr Fau - Brambilla G, Brambilla G Fau - Chiesa G, Chiesa G Fau - Ceresoli G, Ceresoli G Fau - Passera E, et al. A randomized study of lung cancer screening with spiral computed tomography: three-year results from the DANTE trial. (1535-4970 (Electronic)).
13. 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. Eur J Cancer Prev. 2012;21(3):308-15.
14. van Iersel CA, de Koning HJ, Draisma G, Mali WP, 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.
15. 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 Med. 2009;361(23):2221-9.

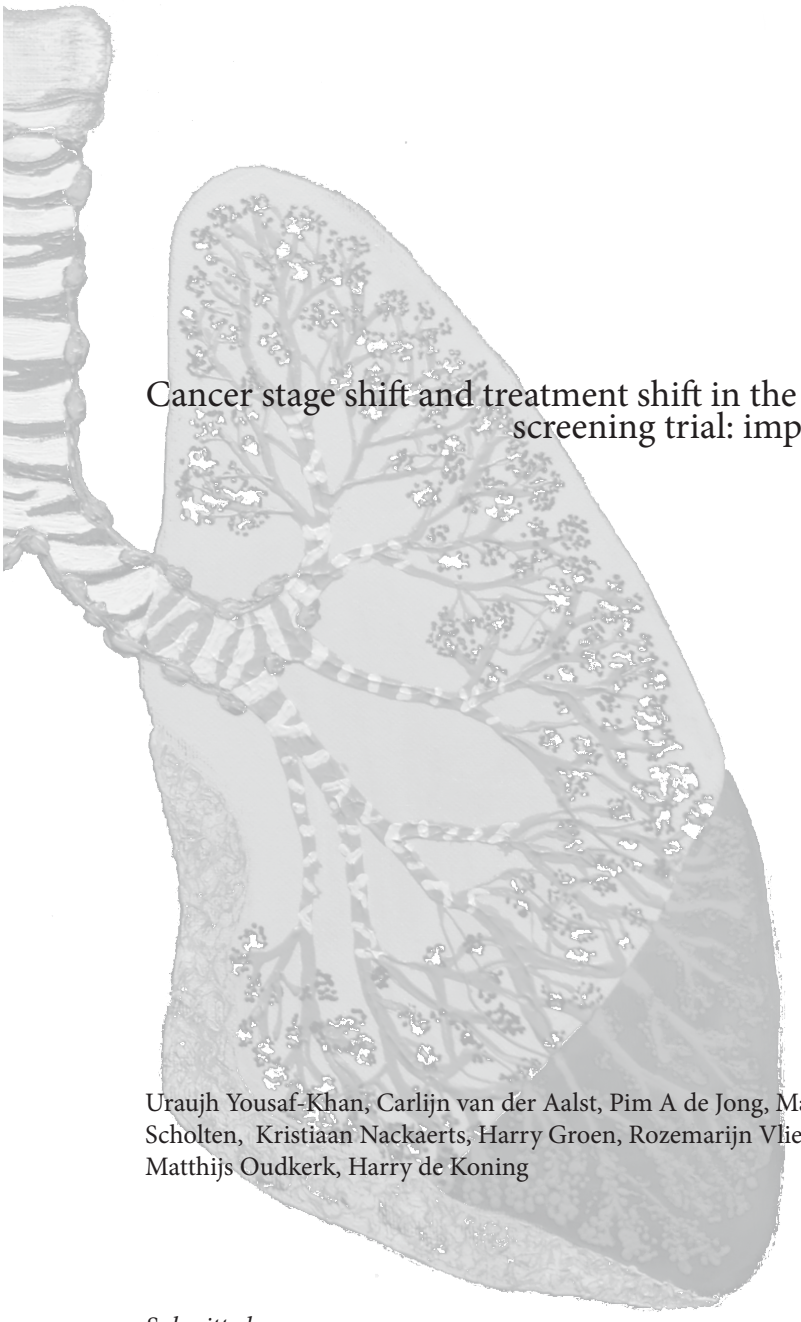
16. Yousaf-Khan U, van der Aalst C, de Jong PA, Heuvelmans M, Scholten E, Lammers JW, et al. Final screening round of the NELSON lung cancer screening trial: the effect of a 2.5-year screening interval. Thorax. 2017;72(1):48-56.
17. Aberle DR, DeMello S Fau - Berg CD, Berg Cd Fau - Black WC, Black Wc Fau - Brewer B, Brewer B Fau - Church TR, Church Tr Fau - Clingan KL, et al. Results of the two incidence screenings in the National Lung Screening Trial. (1533-4406 (Electronic)).
18. Aberle Dr Fau - Berg CD, Berg Cd Fau - Black WC, Black Wc Fau - Church TR, Church Tr Fau - Fagerstrom RM, Fagerstrom Rm Fau - Galen B, Galen B Fau - Gareen IF, et al. The National Lung Screening Trial: overview and study design. (1527-1315 (Electronic)).
19. Cassidy A, Myles Jp Fau - van Tongeren M, van Tongeren M Fau - Page RD, Page Rd Fau - Liloglou T, Liloglou T Fau - Duffy SW, Duffy Sw Fau - Field JK, et al. The LLP risk model: an individual risk prediction model for lung cancer. (0007-0920 (Print)).
20. Tammemagi CM, Pinsky Pf Fau - Caporaso NE, Caporaso Ne Fau - Kvale PA, Kvale Pa Fau - Hocking WG, Hocking Wg Fau - Church TR, Church Tr Fau - Riley TL, et al. Lung cancer risk prediction: Prostate, Lung, Colorectal And Ovarian Cancer Screening Trial models and validation. (1460-2105 (Electronic)).
21. Tammemagi MC, Katki Ha Fau - Hocking WG, Hocking Wg Fau - Church TR, Church Tr Fau - Caporaso N, Caporaso N Fau - Kvale PA, Kvale Pa Fau - Chaturvedi AK, et al. Selection criteria for lung-cancer screening. (1533-4406 (Electronic)).
22. Maisonneuve P, Bagnardi V Fau - Bellomi M, Bellomi M Fau - Spaggiari L, Spaggiari L Fau - Pelosi G, Pelosi G Fau - Rampinelli C, Rampinelli C Fau - Bertolotti R, et al. Lung cancer risk prediction to select smokers for screening CT--a model based on the Italian COSMOS trial. (1940-6215 (Electronic)).
23. Horeweg N, van der Aalst CM, Vliegenthart R, Zhao Y, Xie X, Scholten ET, et al. Volumetric computed tomography screening for lung cancer: three rounds of the NELSON trial. Eur Respir J. 2013;42(6):1659-67.
24. Walter JE, Heuvelmans MA, de Jong PA, Vliegenthart R, van Ooijen PM, Peters RB, et al. Occurrence and lung cancer probability of new solid nodules at incidence screening with low-dose CT: analysis of data from the randomised, controlled NELSON trial. LID - S1470-2045(16)30069-9 [pii] LID - 10.1016/S1470-2045(16)30069-9 [doi]. (1474-5488 (Electronic)).
25. Xu DM, Gietema H, de Koning H, Vernhout R, Nackaerts K, Prokop M, et al. Nodule management protocol of the NELSON randomised lung cancer screening trial. Lung Cancer. 2006;54(2):177-84.
26. Wang Y, van Klaveren RJ, van der Zaag-Loonen HJ, de Bock GH, Gietema HA, Xu DM, et al. Effect of nodule characteristics on variability of semiautomated volume measurements in pulmonary nodules detected in a lung cancer screening program. Radiology. 2008;248(2):625-31.
27. Xie X, Zhao Y Fau - Snijder RA, Snijder Ra Fau - van Ooijen PMA, van Ooijen Pm Fau - de Jong PA, de Jong Pa Fau - Oudkerk M, Oudkerk M Fau - de Bock GH, et al. Sensitivity and accuracy of volumetry of pulmonary nodules on low-dose 16- and 64-row multi-detector CT: an anthropomorphic phantom study. (1432-1084 (Electronic)).

28. Horeweg N, van Rosmalen J, Heuvelmans MA, van der Aalst CM, Vliegenthart R, Scholten ET, et al. Lung cancer probability in patients with CT-detected pulmonary nodules: a prespecified analysis of data from the NELSON trial of low-dose CT screening. Lancet Oncol. 2014;15(12):1332-41.

Part II

Interim stage shift results in the NELSON trial





CHAPTER IV

Cancer stage shift and treatment shift in the NELSON lung cancer screening trial: implications for clinicians

Uraujh Yousaf-Khan, Carlijn van der Aalst, Pim A de Jong, Marjolein Heuvelmans, Ernst Scholten, Kristiaan Nackaerts, Harry Groen, Rozemarijn Vliegenthart, Joachim Aerts, Matthijs Oudkerk, Harry de Koning

Submitted

ABSTRACT

Background
Debate about implementing lung cancer screening in Europe is ongoing. In this study, the cancer stage and treatment shift in the largest European trial, the Dutch-Belgian Lung Cancer Screening Trial (NELSON), were investigated. Furthermore, the generalizability of the trial was assessed.

Method
15,792 participants were evenly randomized into a screen group (7,900) with screening offered at year 1, 2, 4 and 6.5, and into a control group (7,892) that received no screening. The histopathological and treatment characteristics of the first 100 lung cancers diagnosed in each study group were compared. Next, the epidemiological, histopathological and treatment characteristics of the male NELSON control group lung cancers (n=86) were compared with the lung cancers diagnosed in the comparable Dutch birth cohort (1928 to 1953, n= 20,884).

Results
Screen group lung cancers were significantly more often diagnosed in stage I (59.0% vs. 19.0%) and less in stage IV (6.0% vs. 48.0%) compared to the control group ($p<0.001$). Relative to no screening, in the screen group more adenocarcinomas (45.0% vs. 27.0%; $p<0.001$) were observed. Surgical treatment was given more often to the screen group, regardless of the stage, compared to the control group (67.7% vs. 24.5%, $p<0.001$). Lung cancers in the NELSON male control group were diagnosed at a slightly earlier stage compared to Dutch male comparable birth cohort lung cancer patients (stage I or stage II, 29.1% vs. 32.5%; $p<0.001$).

Conclusions
CT screening in the NELSON trial has led to a substantial shift in cancer stage at time of diagnosis and it has led to more treatment to be curative (mainly through surgery). As the actual difference in stage between the NELSON control group and the Dutch comparable birth cohort is small, it is expected that the NELSON results will be generalizable for the target population.

4.1. INTRODUCTION

Lung cancer is a major public health problem world-wide [1, 2]. Despite advances in therapy, the five-year survival for clinically detected lung cancer is still at 16-18%[1]. Smoking is the most important risk factor for lung cancer and smoking cessation is considered to be the most effective strategy to reduce the risk of lung cancer [3]. However, the prevalence of smoking is still high, and even increasing world-wide [4]. Moreover, currently half of the lung cancers are being diagnosed in former smokers [5, 6]. This highlights the need for other additional prevention methods. The National Lung Screening Trial (NLST) presented with the historical results that Computed Tomography (CT) screening and subsequent treatment for lung cancer in high-risk subjects can reduce lung cancer mortality with 20% [7]. Based upon these results, the United States Preventive Screening Task Force (USPSTF) requested an independent review and a modelling study to investigate the long-term harms and benefits of different screening policies. At this moment, the USPTFS recommends annual screening for lung cancer with CT in persons aged 55 through 80 years, with a smoking history of 30 pack-years and currently smoke or have quit within the 15 past years [8, 9]. Although CT screening for lung cancer has been implemented in the United States, the debate about implementing it in Europe is ongoing [10].

In Europe, thus far none of the lung cancer screening trials have shown mortality reduction [11-13]. However, none of these trials were sufficiently powered (in terms of sample size) to identify a significant mortality benefit. The largest European trial, the Dutch-Belgian Lung Cancer Screening Trial (NELSON) has sufficient power to investigate whether CT screening leads to a lung cancer mortality reduction of at least 25% after ten years of follow-up compared to no screening [14]. The NELSON trial is currently in the stage prior to the mortality analysis [15, 16]. A possible mortality benefit will depend on factors such as whether more early stage and less advanced stage lung cancers are diagnosed in the screen arm (the so called cancer stage shift), and whether the early detection enhanced the chances of curative treatment from disease. In the NELSON trial, 69.4% of the screening-detected lung cancers are detected in stage I and 9.8% in stage IIIb/IV [14, 17, 18]. This stage distribution appears to be favorable compared with that of the National Lung Screening Trial (61.6% and 20.0%, respectively) [7, 19]. However, to detect a 'true cancer stage shift', a comparison of all the screen group lung cancers regarding the cancer stage should be made with those lung cancers observed in the control group that did not receive any form of screening.

The aims of this study are (1) to investigate the characteristics (histopathological and treatment) of the lung cancers diagnosed in the control group, (2) to compare them with those lung cancers detected in the screen group, and (3) to assess the generalizability of the results of the NELSON trial by comparing the lung cancers diagnosed in the control group with the lung cancers diagnosed in the Dutch population.

4.2. METHODS

NELSON study

In brief, the NELSON is a population-based, randomized-controlled lung cancer screening trial [15]. Recruitment took place between September 2003 and the second half of 2006, in which addresses of subjects aged between 50 and 74 years were obtained from the population registries of seven districts in the Netherlands and 14 municipalities in Belgium [16]. Eligible where those aged between 50 and 74 years, who had smoked more than 15 cigarettes per day for more than 25 years or more than 10 cigarettes per day for more than 30 years, and were still smoking or quit smoking less than or equal to 10 years ago. Respondents who met the criteria for eligibility received an invitation to participate and an informed consent form. A total of 15,792 persons were randomized (1:1) in a screening group (n=7,900) and in a control group (n=7,892) [20]. Screening took place at year one, two, four and six and a half. The control group received usual care (no screening).

Nodule management strategy

In short, the screening protocol was based on the assessment of volume, nodule growth (defined as change in volume of ≥25%) and volume doubling time (VDT) of the nodules [15, 21, 22]. Results were defined as: 1) negative (new nodule <50mm3 or a previous detected nodule with growth <25% and VDT >600 days); 2) indeterminate (new nodules 50-500mm3 or previously detected nodule with growth ≥25% and VDT of 400-600 days); and 3) positive (new nodules >500mm3 or previously detected nodule with growth ≥25% and VDT <400 days). Those with a negative screening result were screened at the next screening round. Those with an indeterminate result were referred for a short-term follow-up scan (depending on the protocol between 6 weeks and 3 months) until the screening outcome was negative or positive. Those with a positive result were referred to a pulmonologist for a diagnostic work-up and treatment according to the Dutch National guidelines. If their diagnostic work-up was truly positive, the medical data was collected prospectively. If there was no lung cancer, the participants were referred to the next regular screening round.

Study participants

For this sub study, from both study groups the first 100 lung cancer diagnoses were selected [14, 18]. In case of the screen group as well as Dutch as Belgian lung cancer patients were included. In case of the control group, only Dutch lung cancer patients were included as no information at time of analyses from the Belgian control group was available. The histopathological and treatment data of these participants were retrieved from their medical files.

For the comparison between the lung cancers diagnosed in the Dutch NELSON control group and lung cancers diagnosed in the Dutch population, a selection of the Dutch Cancer Society (NKR) was made of subjects belonging to the same birth cohort as the NELSON

trial (1928 to 1953) and with a lung cancer diagnosis between 2005 and 2008 (comparable to the time frame of the NELSON trial). The NKR provided data about the epidemiological characteristics (year of birth, year of diagnosis, age at diagnosis, time between diagnosis and first treatment (in years), time between the first treatment and sub sequent treatments until treatment number eight (in years), vital status at 1st January 2014, year of death and follow-up time (in days)), histopathological characteristics (base for the diagnosis, localization, morphology, differentiation grade, clinical and pathological TNM, first three locations of the metastasis, numbers of investigated lymph nodes, numbers of positive lymph nodes, whether mediastinoscopy was performed and the outcome) and treatment characteristics (which treatment was given (the first eight treatments), and the complications after lung cancer surgery) of the lung cancers diagnosed in the general population as well as in the NELSON trial.

In all cases, the histological features of the lung cancers were coded according to the International Classifications of Diseases, 10th edition (ICD-10). In the NELSON trial the disease stage was determined by using the seventh edition of the TNM staging.

Definitions

Current smokers were those who had smoked in the last seven days before completing the general health questionnaire and those who were still smoking. Former smokers were those who have quit smoking for 10 years or less. Pack-years were defined as 20 cigarettes smoked per day for 1 year. Lung cancers diagnosed by a pulmonologist within 12 to 24 months (depending on the screening round) after a positive screening were defined as screen-detected lung cancers. Interval cancers were lung cancers diagnosed in the screen group after a negative screening and before the next screening round.

Statistical analysis

Continuous variables were tested for normality by using the Kolmogorov-Smirnov test and examining the Q-Q plots. In case the tested variables were distributed normally, the mean and 95% confidence interval (CI) was used to describe the variables. In case the tested continues variables were distributed non-normally, they were described by using medians and Inter Quartile Ranges (IQRs). Differences between the continuous variables were tested by using a Mann-Whitney U test or by analysis of variance. Differences between the categorical variables were tested by using a Mann-Whitney U test or Kruskal-Wallis test. For nominal variables, a $\chi 2$-test or the Fisher's exact test were used. The associations of gender, age at randomization, pack-years smoked and smoking status at randomization with the stage of disease and histology of the lung cancers were assessed through univariate analyses for which logistic regression was used. For all analysis, a p-value of <0.05 was considered as statistically significant, and IBM SPSS Statistics V.21 was used for all analyses.

4.3. RESULTS

Comparison of the lung cancers diagnosed in the screen group and control group

The lung cancers in the screen group were on average diagnosed a year earlier than the lung cancers diagnosed in the control group (2005 vs. 2006, p<0.001; **Table 1**). No differences were observed in age at diagnosis, proportion of males, proportion of current smokers at randomization or smoked pack-years at time of randomization between the two study groups with lung cancer.

Table 1: baseline characteristics of the first 100 control group and screen group participants with lung cancer

	Control group participants with lung cancer		**Screen group participants with lung cancer**		**P-value**
	n	%	n	%	
Total participants	98	-	96	-	-
Total lung cancer	100	-	100	-	-
Males	84	84.0	80	80.0	0.69
Age at diagnosis (median, IQR)	64.5	10.2	65.7	10.7	0.89
Year diagnosis (median, IQR)	2006	1.0	2005	1.0	<0.001
Current smokers[1]	63	64.3	54	56.2	0.25
Smocked pack-years[1] (median, IQR)	43.2	21.0	43.7	24.8	0.17

[1]at time of randomization

Relative to control group lung cancers, significantly more lung cancers in the screen group were diagnosed at stage I (59.0% vs. 19.0%) and less in stage IV (6.0% vs. 48.0%; **Table 2**). Also, the proportion of stage III cancers differed substantially between the two study groups (14.0% vs. 23.0%, respectively for the screen and control group). A higher proportion of adenocarcinomas (45.0% vs. 27.0%) and bronchoalveolar carcinomas (BACs; 3.0% vs. 0.0%) were observed in the screen group compared to the control group, respectively (p<0.001). In the control group, a significantly higher proportion of squamous cell carcinomas (SQMs; 25.0% vs. 21.0%), large cell carcinomas (LCCs; 19.0% vs. 11.0%) and small cell carcinomas (SCLCs; 18.0% vs. 2.0%) was observed. All SCLCs in the control group were detected in stage 4, while in the screen group 11.2% were detected in stage 2, 16.7% in stage 3 and 72.2% in stage 4.

The number of smoked pack-years was the only significant predictor for adenocarcinoma in univariate analysis (p=0.03): those who smoked more pack-years had a higher chance for an

adenocarcinoma lung cancer diagnosis. In case of LCCs only the smoking status significantly predicted lung cancer diagnosis in univariate analysis (p<0.001); those who were current smokers at time of randomization had a higher risk for a LCC diagnosis, compared to the former smokers. In case of SQMs, current smokers had borderline significantly a higher risk for a SQM diagnosis than former smokers (p=0.05). In SCLCs and BACs none of the following variables predicted the outcome: gender (p=0.99 and p=0.39), current smokers (p=0.19 and p=0.85), age at diagnosis (p=0.68 and p=0.68) and smoked pack-years (p=0.45 and p=0.86), respectively.

Table 2: tumor characteristics of the first 100 lung cancers in the control and screen group

		Control group lung cancers		Screen group lung cancers		P-value
		n	%	n	%	
Stage at diagnosis						<0.001
	Ia	9	9.0	55	55.0	
	Ib	10	10.0	5	5.0	
	IIa	8	8.0	11	11.0	
	IIb	2	2.0	-	-	
	IIIa	17	17.0	15	15.0	
	IIIb	6	6.0	8	8.0	
	IV	48	48.0	6	6.0	
	Total	100	100.0	100	100.0	
Histology						<0.001
	Adenocarcinoma	27	27.0	45	45.0	
	SQM	25	25.0	21	21.0	
	BAC	-	-	3	3.0	
	Large cell carcinoma	19	19.0	11	11.0	
	SCLC	18	18.0	2	2.0	
	No diagnosis possible	4	4.0	5	5.0	
	NSCLC, not specified	3	3.0	2	2.0	
	Others	4[1]	4.0	11[2]	11.0	
	Total	100	100.0	100	100.0	
Localization						0.1
	Left	38	38.0	30	30.0	
	Right	59	59.0	69	69.0	
	Both sides	3	3.0	-	-	
	Others	-	-	1[3]	1.0	
	Total	100	100.	100	100.0	

SQM: squamous cell carcinoma; BAC: bronchoalveolar carcinoma; SCLC: small cell carcinoma; NSCLC: non-small cell carcinoma

[1]Others were: 3 large-cell neuro-endocrine tumor and 1 small cell neuro-endocrine tumor; [2] others were: 3 adenosquamous, 3 carcinoid, 1 mixed SCLC/NSCLC, 3 large cell neuro-endocrine tumor, 1 pleiomorph carcinoma; [3] 1 screen-detected lung cancer was located in the carina

Stage I cancers stratified for each study group showed that the proportion of stage Ia cancers was the highest for the screen group (91.7% vs. 47.4%, p <0.001) compared to the control group. The histology was comparable between the two study groups with stage I cancers (p=0.13; **Table 3**).

Screen group lung cancer patients received more often a (potentially curative) surgical treatment regardless of the stage (67.7% vs. 24.5%, p<0.001). Even if stratified for stage Ia, screen group participants more often received a (potentially curative) surgical treatment than control group participants (82.7% vs. 50.0%, p=0.06; data not shown). No difference was seen in age at diagnosis, smoking status and co-morbidity between the screen group and control group participants with stage Ia cancer.

Comparison of the NELSON control group lung cancers versus the lung cancers diagnosed in the comparable Dutch birth cohort

Between 2005 and 2008 in the same birth cohort 32,132 lung cancers were diagnosed. Of this group, 65.0% were males and 35.0% were females. Due to risk-based selection in the NELSON trial, the proportion of female participants was 14.0%. Therefore, further analyses were restricted to male lung cancer patients only.

Lung cancer patients in the comparable Dutch male birth cohort were diagnosed with lung cancer approximately one year later than the NELSON control group lung cancer patients (2007.0 vs. 2006.0, p=0.06; **Table 4**). Compared to the same Dutch birth cohort with lung cancer, the lung cancers diagnosed in the control group were diagnosed in a slightly earlier stage (stage I or stage II, 29.1% vs. 32.5%, p<0.001; **Table 5**). Age at time of diagnosis was not a significant predictor for the stage of lung cancer at time of diagnosis (p=0.38, data not shown).

No significant difference was observed in the first treatment given to a lung cancer patient between the general Dutch population and the NELSON control group (p=0.97); surgery was given in 23.2% vs. 22.5%, and chemotherapy in 50.2% vs. 53.5% respectively.

Table 3: histology and cancer stage distribution of stage Ia en Ib lung cancers, by study group of the NELSON trial

Stage at diagnosis		**Control group lung cancers**		**Screen-detected lung cancers**		**P-value[1]**
		n	%	n	%	<0.001[3]
Ia						
	Adenocarcinoma	3	33.3	27[2]	49.1	0.29[4]
	SQM	3	33.3	10	18.2	
	BAC	-	-	3	5.5	
	Large cell Carcinoma	-	-	4	7.3	
	SCLC	-	-	-	-	
	No diagnosis possible	3	33.3	5	9.1	
	NSCLC, not Specified	-	-	-	-	
	Others	-	-	6	10.9	
	Total	9	100.0	55	100.0	
Ib						0.17[5]
	Adenocarcinoma	4	40.0	3	60.0	
	SQM	5	50.0	-	-	
	BAC	-	-	-	-	
	Large cell Carcinoma	1	10.0	1	20.0	
	SCLC	-	-	-	-	
	No diagnosis possible	-	-	-	-	
	NSCLC, not Specified	-	-	-	-	
	Others	-	-	1	20.0	
	Total	10	100.0	4	100.0	

[1]×-test

[2]Screen group includes one interval cancer

[3]Difference in cancer stage between the two study groups

[4]Difference in histology between two study groups with stage Ia cancer

[5]Difference in histology between two study groups with stage Ib cancer

Table 4: baseline characteristics of the NELSON control group participants with lung cancer and of the comparable Dutch birth cohort with lung cancer (diagnosed between 2005 and 2008)

	NELSON control group lung cancers		**General Dutch population lung cancers**		**p-value**
	n	%	n	%	
Total participants	84	-	20,646	-	
Total lung cancer	86	-	20,884	-	
Age at diagnosis, (median, IQR)	65.4	11.4	68.0	11.0	0.26
Year diagnosis, (median, IQR)	2006	1.0	2007.0	2.0	0.06

Table 5: tumor characteristics of the male lung cancers in the NELSON control group and male lung cancer patients from the Dutch comparable birth cohort

		Control group lung cancers		**Dutch comparable birth cohort lung cancers**		**P-value**
		n	%	n	%	
Stage at diagnosis						<0.001
	Ia	8	9.3	1,379	7.2	
	Ib	9	10.5	1,805	9.4	
	IIa	6	7.0	166	0.9	
	IIb	2	2.3	964	5.0	
	IIIa	14	16.3	2,254	11.8	
	IIIb	5	5.8	3,513	18.4	
	IV	42	48.8	9,043	47.3	
	Total	86	100.0	19,124[1]	100.0	
Histology						0.04
	Adenocarcinoma	24	27.9	4993	24.0	
	SQM	22	25.6	5643	27.1	
	BAC	-	-	340	1.6	
	Large cell carcinoma	14	16.3	3100	14.9	
	SCLC	15	17.4	3031	14.5	
	No diagnosis possible	4	4.7	1636	7.8	
	NSCLC, not specified	3	3.5	1705	8.2	
	Others	4	4.7	396	1.9	
	Total	86	100.0	20,844	100.0	

[1]In 1,760 cases the stage at time of diagnosis was missing

4.4. DISCUSSION

The study shows that screening has not only led to a stage shift in terms of early detection of lung cancer (stage I lung cancers increased by 41.0%), but also to less advanced staged (stage IV) lung cancers by 42.0%, compared to no screening. Furthermore, the proportion of stage Ia cancers are significantly the highest among the screen group lung cancers, compared to the control group lung cancers (91.7% vs. 8.3%). Comparison with the NLST shows, that the stage shift in the NELSON is substantially more favorable: more stage I cancers (60.0% vs. 50.0%), and less stage IV cancers (6.0% vs. 21.7%) were detected in the NELSON trial [7, 14, 17-19]. The NELSON trial used the 7th TNM staging system, whereas the NLST used the 6th TNM staging system, which might explain part of this difference [23]. On the other hand, in the NELSON trial fewer females were randomized (due to study selection criteria), compared to the NLST (16.5% vs. 41.0%), while females are diagnosed with lung cancer at a lower stage [24]. Furthermore, the NELSON screening protocol consisted of increasing screening intervals between the four screening rounds (1, 2, and 2.5 years), whereas the NLST screening protocol consisted of three annual screening rounds [15, 25]. However, the current NELSON analysis includes only the first 100 lung cancer's diagnosis of each study group. Thereby, the second screening round was still running in the screen group. Altogether, this sub study showed that it seems that the NELSON screening strategy is at least as capable as the NLST to diagnose lung cancer at a more favorable stage compared to the current setting (no screening) in the Netherlands and Belgium. In general, compared to the European trials, the NELSON trial shows a more favorable stage shift [12, 13, 26]. This may be explained by the substantially larger sample size and a different screening strategy. The DLCST used an adapted version of the NELSON screening strategy [11]. Relative to the Danish Lung Cancer Screening Trial (DLCST), no difference in stage is observed.

Inherent to a stage shift is the so called "histology shift". Debate is ongoing whether the early detection of lung cancers is based on the detection of lung cancer in an early stage, or rather an increased detection of slow-growing tumors (e.g. adenocarcinomas and BACs) [27-29]. Several studies have reported that up to 25% of the CT detected lung cancers are relatively slow-growing, and that up to 80% of such indolent cancers are adenocarcinomas or BACs [29-32]. In the current NELSON analysis, 48.0% of the screen group lung cancers are relatively slow-growing tumors, 93.8% of which are adenocarcinomas. Detection of more relatively slow-growing tumors may lead to weakening the effect of screening and may result in unnecessary diagnosis and over-treatment [27, 33]. However, in the NELSON trial less female participants were included. It is known that adenocarcinomas are the most prevalent tumor histology among females [24]. Furthermore, screening has also led to more SCLCs to be diagnosed at a lower stage compared to no screening. Further studies, for example by use of Microsimulation Screening Analysis (MISCAN) modelling, are needed to investigate the histology shift in more depth and to determine the proportion of overdiagnosed lung cancers in the NELSON trial.

A favorable cancer stage shift is important, as it is know that early detection of lung cancer is related to the possibility of surgical resection and a better survival [2, 7, 8, 30]. In the NELSON

trial, the screen group lung cancer patients received a surgical treatment 2.8 times more often regardless of stage at time of diagnosis than the control group lung cancer patients. Almost the same portion of screen group lung cancer patients were treated surgically (alone or with chemotherapy and/or radiotherapy) as in the NLST [7]. Compared to the ITALUNG trial less screen-detected lung cancers were treated surgically in the NELSON trial (67.7%% vs. 85.4%) [34]. However, in comparison with the current NELSON analysis, the sample size in ITALUNG was 2.4 times smaller, and therefore a comparison should be made with precaution. Nevertheless, implementation of lung cancer screening will probably lead to more surgical treatment options of lung cancer. And it will open a new area of investigations in surgical resections of small sized tumors.

Furthermore, this study shows that screen group stage Ia cancer patients received a surgical treatment slightly more often than the stage Ia control group patients. Clinically, most lung cancers are diagnosed in an advanced disease stage [2]. To be diagnosed with a stage I cancer as a control group participant, the control group participant should typically have been under the supervision of a specialist for any reason (e.g. COPD treatment). In this sub study, no difference was observed in age at diagnosis or in co-morbidity (e.g. COPD) between both study group subjects with a stage Ia cancer. It is unknown what caused the difference in treatment in the stage I lung cancers.

Additionally, in interpreting the mortality results of a screening trial, it is also necessary to know whether the study population was representative of the target population. Previous publications reveal that the NELSON control group participants were slightly younger, healthier, higher educated and more willing to participate in a screening program, compared to those who were eligible to participate but did not [35]. This led to modest differences in mortality compared to the eligible non-responders. The observed differences were small and it is assumed that it is unlikely that it will influence the generalizability of the NELSON trial. However, it was unknown until then whether the control group lung cancers were representative of the lung cancers observed in the general population. Current study shows that the lung cancers outside the NELSON are diagnosed slightly in a higher stage than in the NELSON control group. One potential explanation is the study eligibility in which subjects with a history of cancer (lung cancer in the past five years or longer than five years ago but still under treatment, or those with renal cancer, melanoma or breast cancer) or those with a moderate or bad self-reported health who were unable to climb two flight of stairs and those with a CT chest examination less than one year ago before they filled in the NELSON questionnaire, were excluded from participation in the NELSON trial [15, 16]. As the actual difference in stage between the NELSON control group and the Dutch comparable birth cohort is small, it is expected that the NELSON results will still be generalizable for the target population.

Major strengths of this study are the population-based recruitment and verified lung cancer diagnosis by obtaining medical records of each lung cancer patient. Limitation of the current NELSON analysis is that at the moment it is not possible to investigate the correlation

between stage shift, treatment shift and lung cancer mortality (reduction). Furthermore, there is the obstacle of the lack of data about the general health and smoking related health of the comparable Dutch birth cohort.

In conclusion, current NELSON analysis shows that LDCT screening has so far led to a substantial favorable shift in cancer stage at time of diagnosis. Implementation of lung cancer screening will lead to a higher rate of early stage lung cancers, and subsequently a shift in treatment options from mainly advanced disease stage treatment (e.g. palliative) to curative treatment (mainly surgery).

ACKNOWLEDGMENTS

We thank the Dutch Cancer Registry (NKR) for providing the data about the comparable Dutch cohort with lung cancer and for the data linkages of the NELSON cohort with the registry. We thank the system controllers R. Faber and F.J.P. Santegoets, and the secretary M.Quak (all from the department of Public Health, Erasmus University Medical Center Rotterdam) for their contribution and maintenance of the database. Furthermore, we thank R. Ziengs (University Medical Center Groningen) and S. van Amelsvoort-van der Vorst (University Medical Center Utrecht) for their contribution.

REFERENCES

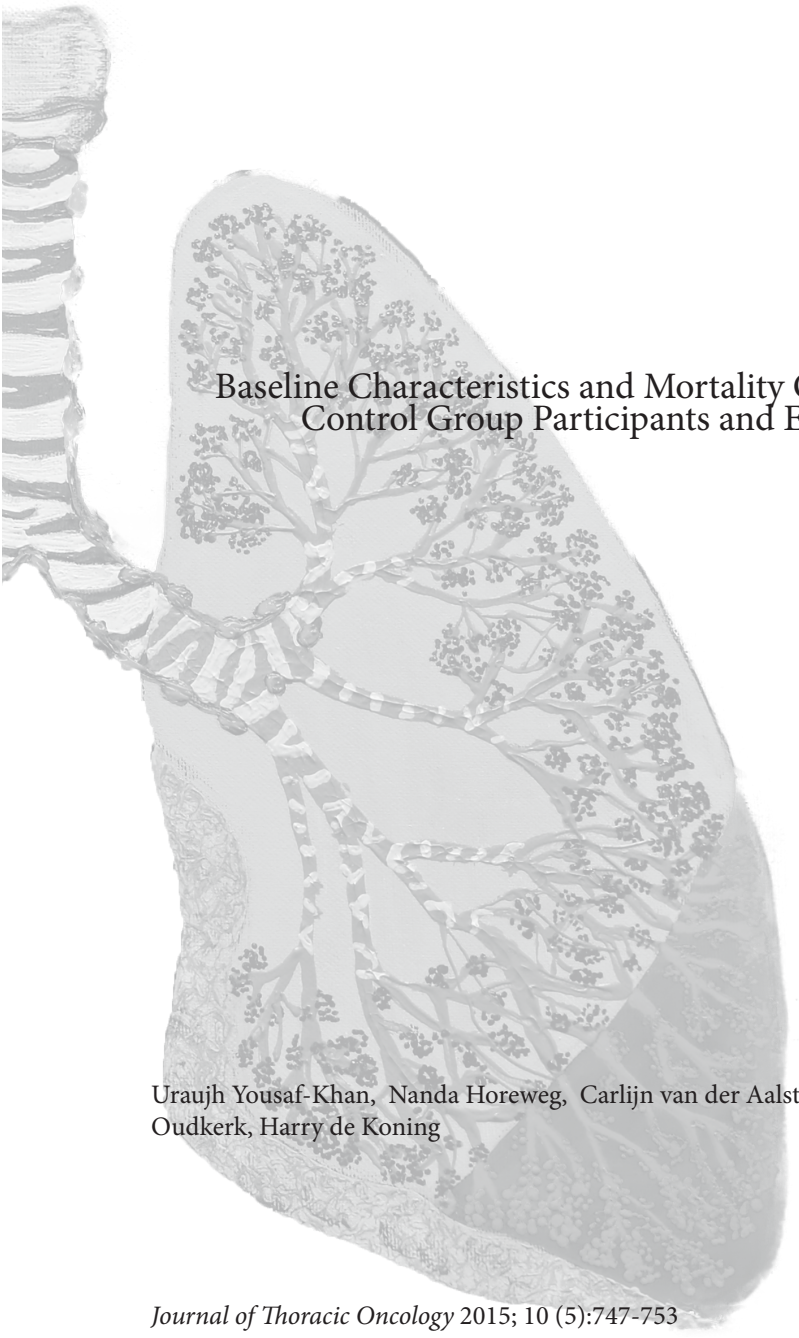
1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. CA Cancer J Clin. 2016;66(1):7-30.
2. Torre LA, Siegel RL, Jemal A. Lung Cancer Statistics. Adv Exp Med Biol. 2016;893(0065-2598 (Print)):1-19.
3. Burns DM. Primary prevention, smoking, and smoking cessation: implications for future trends in lung cancer prevention. Cancer. 2000;89(11 Suppl):2506-9.
4. Ng M, Freeman MK, Fleming TD, Robinson M, Dwyer-Lindgren L, Thomson B, et al. Smoking prevalence and cigarette consumption in 187 countries, 1980-2012. Jama. 2014;311(2):183-92.
5. Tong L, Spitz MR, Fueger JJ, Amos CA. Lung carcinoma in former smokers. Cancer. 1996;78(5):1004-10.
6. Mong C, Garon EB, Fuller C, Mahtabifard A, Mirocha J, Mosenifar Z, et al. High prevalence of lung cancer in a surgical cohort of lung cancer patients a decade after smoking cessation. J Cardiothorac Surg. 2011;6:19.
7. National Lung Screening Trial Research T, 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.
8. Moyer VA, Force USPST. Screening for lung cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2014;160(5):330-8.
9. 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.
10. van der Aalst CM, Ten Haaf K, de Koning HJ. Lung cancer screening: latest developments and unanswered questions. (2213-2619 (Electronic)).
11. 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.
12. 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. Eur J Cancer Prev. 2012;21(3):308-15.
13. Infante M, Cavuto S, Lutman FR, Passera E, Chiarenza M, Chiesa G, et al. Long-Term Follow-up Results of the DANTE Trial, a Randomized Study of Lung Cancer Screening with Spiral Computed Tomography. Am J Respir Crit Care Med. 2015;191(10):1166-75.
14. Yousaf-Khan U, van der Aalst C, de Jong PA, Heuvelmans M, Scholten E, Lammers JW, et al. Final screening round of the NELSON lung cancer screening trial: the effect of a 2.5-year screening interval. Thorax. 2017;72(1):48-56.
15. 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 Med. 2009;361(23):2221-9.
16. van Iersel CA, de Koning HJ, Draisma G, Mali WP, 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.

17. Horeweg N, van der Aalst CM, Vliegenthart R, Zhao Y, Xie X, Scholten ET, et al. Volumetric computed tomography screening for lung cancer: three rounds of the NELSON trial. Eur Respir J. 2013;42(6):1659-67.
18. Horeweg N, van der Aalst CM, Thunnissen E, Nackaerts K, Weenink C, Groen HJ, 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.
19. Aberle DR, DeMello S Fau - Berg CD, Berg Cd Fau - Black WC, Black Wc Fau - Brewer B, Brewer B Fau - Church TR, Church Tr Fau - Clingan KL, et al. Results of the two incidence screenings in the National Lung Screening Trial. (1533-4406 (Electronic)).
20. Yousaf-Khan U, van der Aalst C, de Jong PA, Heuvelmans M, Scholten E, Walter J, et al. Risk stratification based on screening history: the NELSON lung cancer screening study. Thorax. 2017.
21. Wang Y, van Klaveren RJ, van der Zaag-Loonen HJ, de Bock GH, Gietema HA, Xu DM, et al. Effect of nodule characteristics on variability of semiautomated volume measurements in pulmonary nodules detected in a lung cancer screening program. Radiology. 2008;248(2):625-31.
22. Xu DM, Gietema H, de Koning H, Vernhout R, Nackaerts K, Prokop M, et al. Nodule management protocol of the NELSON randomised lung cancer screening trial. Lung Cancer. 2006;54(2):177-84.
23. Marshall HM, Leong SC, Bowman RV, Yang IA, Fong KM. The science behind the 7th edition Tumour, Node, Metastasis staging system for lung cancer. Respirology. 2012;17(2):247-60.
24. Kligerman S, White C. Epidemiology of lung cancer in women: risk factors, survival, and screening. AJR Am J Roentgenol. 2011;196(2):287-95.
25. Aberle Dr Fau - Berg CD, Berg Cd Fau - Black WC, Black Wc Fau - Church TR, Church Tr Fau - Fagerstrom RM, Fagerstrom Rm Fau - Galen B, Galen B Fau - Gareen IF, et al. The National Lung Screening Trial: overview and study design. (1527-1315 (Electronic)).
26. 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.
27. 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.
28. 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.
29. Wang Z, Hu Y, Wang Y, Han W, Wang L, Xue F, et al. Can CT Screening Give Rise to a Beneficial Stage Shift in Lung Cancer Patients? Systematic Review and Meta-Analysis. PLoS One. 2016;11(10):e0164416.
30. Young RP, Hopkins RJ. Mortality Reduction, Overdiagnosis, and the Benefit-to-Harm Ratio of Computed Tomography Screening. Am J Respir Crit Care Med. 2015;192(3):398-9.
31. Young RP, Hopkins RJ. Stage shift in computed tomography screening: possible role of indolent cancers, "histology shift," and overdiagnosis. Am J Respir Crit Care Med. 2013;188(8):1034-5.

32. Wilson DO, Ryan A, Fuhrman C, Schuchert M, Shapiro S, Siegfried JM, et al. Doubling times and CT screen-detected lung cancers in the Pittsburgh Lung Screening Study. Am J Respir Crit Care Med. 2012;185(1):85-9.
33. Han SS, Ten Haaf K, Hazelton WD, Munshi VN, Jeon J, Erdogan SA, et al. The impact of overdiagnosis on the selection of efficient lung cancer screening strategies. Int J Cancer. 2017;140(11):2436-43.
34. 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. J Thorac Oncol. 2013;8(7):866-75.
35. Yousaf-Khan U, Horeweg N, van der Aalst C, Ten Haaf K, Oudkerk M, de Koning H. Baseline Characteristics and Mortality Outcomes of Control Group Participants and Eligible Non-Responders in the NELSON Lung Cancer Screening Study. J Thorac Oncol. 2015;10(5):747-53.

Part III

the cause of death of the NELSON study participants



CHAPTER V

Baseline Characteristics and Mortality Outcomes of NELSON Control Group Participants and Eligible Nonresponders

Uraujh Yousaf-Khan, Nanda Horeweg, Carlijn van der Aalst, Kevin ten Haaf, Mathijs Oudkerk, Harry de Koning

Journal of Thoracic Oncology 2015; 10 (5):747-753

ABSTRACT

Background
Individuals who are younger have a high socioeconomic background and/or have a healthy lifestyle are more inclined to participate in screening trials. This form of bias may affect the generalizability of study results to the target population. This study aimed to investigate the generalizability of the NELSON lung cancer screening trial to the Dutch population.

Methods
People at high risk for developing lung cancer were identified by sending a health questionnaire to 606,409 persons aged 50–74 years, based on population registries. Eligible subjects received an invitation to participate (n = 30,051). 15,822 subjects agreed to participate and were randomized, whereas 15,137 did not respond (so called eligible nonresponders). Baseline characteristics and mortality profiles were compared between control group participants and eligible nonresponders.

Results
Participants had better self-reported health (p=0.02), were younger, more physically active, higher educated, and more often former smokers compared with eligible nonresponders (all p<0.001). No differences were seen in self-reported outcomes of pulmonary tests, history of lung cancer, and smoked pack-years. Mortality due to all-causes (p<0.001) and mortality classification separately was lower among participants. However, the proportion of subjects death due to cancer was higher among participants (62.4% vs. 54.9%).

Conclusions
Modest differences in baseline characteristics between participants and eligible nonresponders, led to minor differences in mortality profiles. However, group sizes were large and therefore it seems unlikely that these small differences will influence the generalizability of the NELSON trial. Results of the NELSON trial can roughly be used to predict the effect of population-based lung cancer screening.

5.1. INTRODUCTION

Lung cancer is a major public health problem worldwide, due to its high incidence and poor 5-year survival rate of less than 15% [1]. Smoking cessation offers the best prospects for reducing the risk of developing lung cancer [2]. Although smoking prevalence is decreased in Europe[3], the residual effects of smoking on lung cancer risk remains notable in former smokers and a significant proportion of lung cancers are now diagnosed in former smokers [4, 5]. For this group, primary prevention is not meaningful. However, if lung cancer is detected in an early stage, treatment options are generally more promising [6]. The National Lung Screening Trial demonstrated that computed tomography (CT) screening can reduce lung cancer mortality by 20% compared with chest radiography [7]. In the United States, this finding has led the United States Prevention Service Task Force to recommend lung cancer screening for current and former smokers, if quit within the past 15 years, aged 55 through 80 years with a smoking history of at least 30 pack-years [8]. However, many issues remain regarding the technical and logistical aspects of screening, cost–effectiveness and generalizability. In Europe, no lung cancer screening trial has yet demonstrated a significant reduction in lung cancer mortality [9-11]. However, the largest European trial, the Dutch-Belgian lung cancer screening trial (NELSON), is still ongoing. The NELSON trial investigates whether screening using LDCT can reduce lung cancer mortality by at least 25% at 10 years of follow-up [12, 13]. Major differences between the NELSON trial and the National Lung Screening Trial are that NELSON (1) offers no screening to control group participants, (2) has different intervals between screening rounds, and (3) uses different management protocols for nodules and abnormalities [12, 14]. In interpreting the results of screening studies, it is important to know whether study participants were representative of the target population, as volunteers who are healthier and more concerned about their own health are more willing to participate in screening programs [15-17]. This form of bias may affect the generalizability of the study results, as the studied subjects may differ from the target population for screening.

So far, previous studies have indicated that participants of lung cancer screening studies are younger, [16, 18, 19] less likely to be current smokers, [15, 17, 19, 20] more physically active and higher educated [15, 18, 19] compared with nonparticipants. Other cancer screening studies indicated that higher socioeconomic status and "healthy lifestyle" predicts screening participation [18, 21-24]. Screening trial participants also had lower incidence of cancer [15, 17], diabetes [15], cardiovascular [15], and respiratory [15] diseases, than nonparticipants. One pilot study of lung cancer screening even showed that participants had a lower mortality rate for all types of cancer besides lung cancer, cardiovascular diseases, and noncancerous diseases other than cardiovascular and respiratory diseases compared with nonparticipants [16]. However, lung cancer mortality was higher among participants. This mortality difference might be explained by selection bias; attendees of screening programs may have more awareness of being at risk of developing lung cancer, which may increase their interest in screening [16, 25].

So far, previous research showed that the NELSON study population is younger, has a better general health, has a higher proportion of current heavy smokers and is slightly lower

educated compared with the general Dutch population [17]. However, less is known about potential differences in physical activity, alcohol consumption, smoking-related symptoms, the effect on the mortality profile of participants and eligible nonparticipants (the so-called eligible nonresponders; Fig. 1). The aim of this study was to investigate whether differences in characteristics and mortality profiles of participants of the NELSON study, and eligible nonresponders exist. The results of this study are relevant for the interpretation of the forthcoming mortality analyses of the NELSON trial.

5.2. METHODS

NELSON Trial

In the NELSON trial, 15,822 high-risk volunteers were randomized (1:1) to screening (n = 7,915) using LDCT at respectively baseline and 1, 3, and 5.5 years after baseline, or to no screening (n = 7,909) [26].The NELSON study aims to investigate whether screening using LDCT can reduce lung cancer mortality by at least 25%. The study design and conduct were published previously [27-29]. The NELSON trial was approved by the Dutch Minister of Health after positive advice from the Dutch health Council and by the Ethics Boards of the participating centers.

Study Population

Population-based recruitments

During the recruitment phase, which occurred in two waves (during the second half of 2003 and the second half of 2005), addresses of subjects aged between 50 and 74 years were obtained from the population registries of seven districts in the Netherlands and 14 municipalities around Leuven in Belgium [28]. These subjects received a questionnaire about their general health, medical check-ups and history, physical activity, body weight and length, smoking history, alcohol consumption, family history of cancer, education and their opinion on screening programs in general.

General health was determined by the subjects "ability to climb two flights of stairs" (yes, no, don't know) and how they would describe their health: excellent, very good, good, moderate, or severe. Questions regarding smoking-related symptoms of lung disease were: did you have symptoms of coughing/sputum/wheezing/dyspnea for at least 3 months this year? (yes, no).

Questions on medical history and check-ups were as follows: was one of the following diagnostic procedures performed last year, 1–5 year, or greater than or equal to 5 years ago: (1) chest x-ray, pulmonary function test, CT-scan of the chest or sputum test?, (2) did you undergo lung surgery (e.g., pneumonectomy or lobectomy)?, (3) were you diagnosed with cancer and if so, when (less than 5 years ago, greater than or equal to 5 years ago, or greater

than or equal to 5 years ago and still under treatment)? and (4) what type of cancer were you diagnosed with (lung cancer, breast cancer, kidney cancer, melanoma, or other type)? Furthermore, physical activity was assessed as follows: how many times a week are you physically active for greater than or equal to 30 minutes (daily, 5–6 times, 2–4 times, 1 time, or less than 1 time a week). Alcohol consumption was assessed by asking how much alcohol was consumed at once (in pints) and at which base: daily, 5–6 times a week, 3–4 times a week, 1–2 days a week, 1–3 days a month, less than 1 glass a month or never. Willingness to participate in screening programs was assessed for prostate cancer, colon cancer, diabetes, cholesterol and cardiovascular diseases (yes, no, do not know) and their opinion on an acceptable number of persons to screen to detect one case of lung cancer at early stage (10, 100, 1000, 10,000, 100,000, or 1,000,000). The highest completed level of education was determined through a single question with seven options: primary education, lower technical or vocational education, general secondary education, secondary technical or vocational education, senior general secondary education or pre-university education, higher technical or vocational education and university. The questionnaire also assessed smoking in detail [28, 30]. Finally, each person's body mass index was calculated (weight/length2).

A total of two recruitment rounds were necessary to reach the required number of participants. The questions of the first questionnaire were slightly changed for the second wave using the experience of the first response. The overall response rate for the first questionnaire was 24.9% [12].

Respondents who met the eligibility criteria (n=30,051) received an invitation to participate, a second questionnaire, an information leaflet and informed consent form for the NELSON trial [28]. The eligibility criteria were as follows: age 50–75 years, smoking history of greater than or equal to 15 cigarettes per day for greater than or equal to 25 years or greater than or equal to 10 cigarettes for greater than or equal to 30 years, and were still smoking or had quit less than or equal to 10 years ago. Exclusion criteria were: a moderate or bad self-reported health and inability to climb two flight of stairs, a body weight greater than or equal to 140 kg, a history of renal, melanoma or breast cancer, lung cancer diagnosed less than 5 years ago or greater than 5 years ago but still receiving treatment, or a chest CT examination within the past year [28]. In addition, the second questionnaire assessed smoking habits and exposure to asbestos in more detail.

Eligible responders who provided informed consent and completed the second questionnaire (n=15,822, response rate of 51.1%) were randomized (1:1) to either the screen group or the control group.

Inclusion in this substudy

For this substudy, subjects randomized to the control group (n=7453) were compared with eligible subjects who did not participate (n=13,661). Subjects randomized to the screen group were excluded because of the potential effect of screening on their mortality profiles and the

embargo on mortality outcomes of this group. Furthermore, this substudy was limited to Dutch subjects, as only Dutch mortality data was available at the time of analyses.

Mortality Data

Anonymised mortality data for both groups were obtained via Statistics Netherlands. January 2013 was chosen as end date of this substudy, at which point 99.1% of the subjects were traceable. To obtain mortality data, this study population was matched using four variables: sex, date of birth, zip code, and date of obtaining addresses. This led to an accuracy of almost 98% in matching.

Person-years were calculated as the time between obtaining the addresses of the subject and subject's date of death or the end date of this study, whichever came first.
To analyze mortality profiles, we classified the causes of death by disease groups, using the International Classification of Diseases, 10th edition: all-causes, all cancer causes, cardiovascular diseases, respiratory diseases, and noncancer diseases other than cardiovascular or respiratory diseases.

Statistical Analyses

Baseline characteristics of control group subjects and eligible nonresponders were retrieved from the first questionnaire. Differences in baseline characteristics were assessed using the following tests: for continuous variables, normality was tested using the Kolmogorov–Smirov test and differences between the two groups were assessed by using the Mann– Whitney *U* test, as appropriate. For nominal variables, the $\chi 2$ test was used and the Mann–Whitney *U* test was used for categorical variables.

Classified mortality data were compared between the two groups by using the $\chi 2$ test. For all analyses, *p* values less than 0.05 were considered statistically significant. SPSS version 21 and STATA 13 special edition were used to perform the analyses.

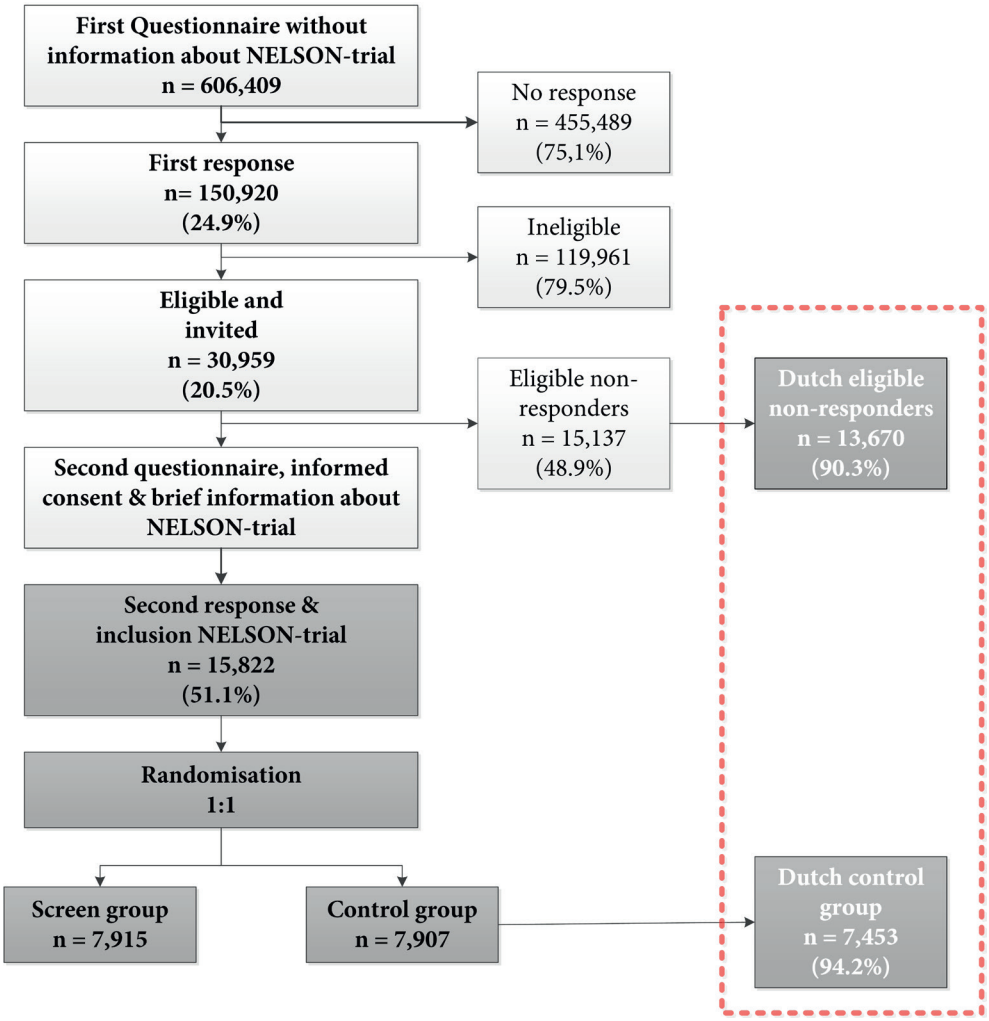


Figure 1. recruitment of the NELSON study participants and the selection for this substudy.

5.3. RESULTS

A total of 7,453 Dutch control group participants were compared with 13,661 Dutch eligible nonresponders (**Table 1**). Participants were younger (p<0.001), more often male (p<0.001), had better self-reported general health (p=0.02), higher level of physical exercise (p<0.001) and a higher level of education (p<0.001). Participants also consumed more alcohol (p<0.001) and consisted of higher proportion of former smokers (p<0.001). However, most differences in proportions were small.

Small differences were also seen in smoking-related characteristics (**Table 2**). Smoking duration was lower among participants (p<0.001), whereas numbers of cigarettes smoked per day was higher among participants (p<0.001). However, no differences were observed in the number of pack-years smoked between participants and eligible nonresponders. Participants started smoking at a younger age (p<0.001) and were more willing to quit smoking than eligible nonresponders (p<0.001). Among current smokers, participants were more often in an advanced stage- according to the stages of change- to quit smoking compared with eligible Nonresponders (p<0.001). Participants reported significantly more smoking-related symptoms (p=0.04) and had undergone a pulmonary function test more often (p<0.001). However, no differences were seen in the self-reported outcome of these pulmonary function tests (p=0.28).

During the study period, the all-cause mortality rate among eligible nonresponders was higher compared with the participants (p<0.001; **Table 3**).The eligible nonresponders had a higher mortality rate due to all types of cancer (p=0.002), cardiovascular diseases (p<0.001), respiratory diseases (p=0.018), and noncancerous diseases other than cardiovascular or respiratory (p<0.001). However, the proportion of deaths due to cancer was higher among participants (62.4% vs. 54.9%). Higher educational achievement was significantly associated with higher mortality from all types of cancer (χ2 17.3; p<0.001). Furthermore, a longer follow-up was seen for participants (10 years vs. 9 years).

Participants were significantly more likely to participate in any of the mentioned screening programs compared with the eligible nonresponders (all p<0.001, data not shown). The median physical distance from home to one of the nearby participating screening centers was significantly less for eligible nonresponders than for participants (16.9 km versus 17.9 km; p=0.003).

Table 1: baseline characteristics of NELSON control group participants and eligible non-responders.

	Control group participants		**Eligible non-responders**		**p-value**
	Total N = 7,453		**Total N = 13,661**		
	%	n/N	%	n/N	
Age (years), median (IQR)	57.0 (8.0)	7,453	58.0 (9.0)	13,661	< 0.001
Male	84.2	6,275/7,453	80.6	11,013/13,661	< 0.001
General health					0.02
Excellent/very Good	15.2	1,124/7,393	14.2	1,913/13,477	
Good	66.6	4,922/7,393	66.7	8,984/13,477	
Moderate/poor	18.2	1,347/7,393	19.1	2,580/13,477	
Physical exercise[a]					< 0.001
High	44.5	3,292/7,398	48.5	6,533/13,459	
Moderate	44.8	3,318/7,398	39.8	5,354/13,459	
Low	10.7	788/7,398	11.7	1,572/13,459	
BMI, median (IQR)	25.9 (4.2)	7,177/7,453	25.8 (4.4)	12,932/13,661	0.16
Education level[b]					< 0.001
Lowest	11.0	806/7,352	18.1	2,410/13,339	
Low	37.4	2,750/7,352	41.4	5,530/13,339	
Medium	23.3	1,712/7,352	20.6	2,750/13,339	
High	28.3	2,084/7,352	19.9	2,649/13,339	
Alcohol[c]**,** median (IQR)	15.7 (83.9)	6,754/7,453	13.8 (83.9)	11,705/13,661	< 0.001
Smoker status					< 0.001
Current smoker	54.8	4,077/7,434	60.4	8196/13,578	
Former smoker	45.2	3,357/7,434	39.6	5382/13,578	
History of lung cancer	4.7	344/7,396	4.4	594/13,502	0.40
Person-years of observation, median (IQR)	10.0 (2.0)	7,453	9.0 (2.0)	13,661	< 0.001

Data were presented as % (*n/N*) unless stated otherwise.

[a]Physical activity: high was defined as greater than or equal to 5 times active for greater than or equal to 30 minutes a week, moderate was defined as greater than or equal to 1 but
less than 5 times active for greater than or equal to 30 minutes a week and low was defined as less than 1 time active for greater than or equal to 30 minutes a week.

[b]Education level: lowest: only elementary; low education: Lower technical or vocational education and general secondary education; medium education level: secondary technical
or vocational education and senior general secondary education; high education level: higher technical or vocational education and university.

[c]Alcohol consumption in glasses per week.

BMI, body mass index.

Table 2: smoking related characteristics of NELSON control group participants and eligible non-responders

	Control group participants		**Eligible non-responders**		**p-value**
	Total N = 7,453		**Total N = 13,661**		
	%	n/N	%	n/N	
Pack-years, median (IQR)	37.9 (19.8)		37.9 (21.5)		0.07
Smoking duration					< 0.001
≤ 35 yrs	38.2	2,838/7,437	35.7	4,859/13,594	
36-40 yrs	31.8	2,363/7,437	30.4	4,136/13,594	
41-45 yrs	19.5	1,451/7,437	20.7	2,816/13,594	
> 45 yrs	10.5	785/7,437	13.1	1,783/13,594	
No. of cigarettes smoked per day					< 0.001
≤ 15	22.1	1,642/7,439	24.2	3,282/13,604	
16-20	28.3	2,101/7,439	29.5	4,016/13,604	
21-25	26.8	1,994/7,439	25.3	3,444/13,604	
26-30	11.0	822/7,439	9.4	1,283/13,604	
31-40	6.9	513/7,439	7.1	965/13,604	
>40	4.9	367/7,439	4.5	614/13,604	
Starting age of smoking					< 0.001
≤ 14 yrs, n (%)	16.0	1,189/7422	15.0	2,000/13,537	
15-19 yrs, n (%)	65.0	4,845/7422	64.0	8,718/13,537	
20-24 yrs, n(%)	16.0	1,184/7422	17.0	2,338/13,537	
>25 yrs, n (%)	3.0	204/7422	4.0	481/13,537	
Motivated to quit smoking[a]	93.3	4,854/5,201	91.4	7,555/8,269	< 0.001
Stage of Change[b]					< 0.001
Precontemplation phase	33.2	1,725/5,201	39.3	3,249/8,269	
Contemplation phase	14.1	736/5,201	13.7	1,129/8,269	
Preparation	7.3	379/5,201	7.5	621/8,269	
Action	4.7	244/5,201	4.3	359/8,269	
Maintenance	40.7	2,117/5,201	35.2	2,911/8,269	
Smoking related symptoms[c]					0.04
Yes, ≥1	53.8	2,777/5,157	52.0	4,279/8,229	
Spirometry					
Yes	59.2	2978/5,029	53.2	4,186/7,873	< 0.001
Result of spirometry[d]					0.28
Normal (%)	81.3	1,867/2,296	82.5	2,522/3,058	
Abnormal (%)	18.7	429/2,296	17.5	536/3,058	

Data were presented as % (*n/N*) unless stated otherwise.

[a]Motivated to quit smoking: comparison between subjects who are current smokers only.

[b]Stage of change: precontemplation phase: does not want to stop, wants to stop but not in the next 5

years, wants to stop but not in the next year, wants to stop but not in the next 6 months. Contemplation phase: wants to stop in the next 6 months. Preparation: wants to stop in the next 1 month. Action: stopped less than 6 months ago. Maintenance: stopped greater than 6 months ago.

[c]Smoking-related symptoms: coughing, sputum, dyspnea, and wheezing.

[d]Spirometry: comparison between subjects with spirometry only.

Table 3: mortality rates (per 1.000 person-years) by causes of NELSON control group participants and eligible non-responders

Cause of death	**Control group**		**Eligible non-responders**		**Mortality rate ratio**	**p-value**
	%	Rate	%	Rate		
All cancer types	62.4	6.32	54.9	7.59	0.83	0.002
Cardiovascular diseases (CVD)	20.2	2.05	24.0	3.32	0.62	<0.001
Respiratory diseases	4.4	0.45	5.3	0.73	0.61	0.018
Noncancerous diseases other than CVD or respiratory diseases	12.9	1.30	15.8	2.19	0.59	<0.001
All causes	9.1	10.11	11.2	13.83	0.73	<0.001

[a]p-value for mortality rate ratio = 1.

CVD, cardiovascular diseases.

5.4. DISCUSSION

This study investigated differences in characteristics and mortality profiles of participants of the NELSON trial and eligible nonresponders. Results of this study are essential to determine whether mortality results of the NELSON trial are generalizable to the target Dutch population for lung cancer screening.

Participants of the NELSON trial were significantly younger, had better self-reported health, were more physically active, and higher educated compared with eligible nonresponders, although the differences in proportions were modest. These results are in line with previous studies in cancer screening trials [15, 16, 18, 19]. Furthermore, men were more likely to participate in the NELSON trial, whereas more women participated in the Danish Lung cancer Screening Trial (DLST) [18].

Different recruitments methods may explain the differences in study populations between NELSON trial and DLST: the NELSON trial was designed to recruit only men at first, because of fewer Dutch women met the smoking-related inclusion criteria of the NELSON study. However, in the second recruitment women were also invited to allow the NELSON study results to be generalizable to women. In contrast, the DLST recruited both sexes from the start of the study. Such overrepresentation of women participating in screening trials is also seen by others and may be because women are more used to screening from other cancer screening programs [31].

In the NELSON study, number of pack-years smoked between the two groups was similar, but participants were more often former smokers [17].This is in contrast with the DLST and an Italian lung cancer screening trial, in which current smokers were overrepresented [16, 18]. However, the DLST also reported that despite active smoking, participants were more willing to quit smoking than nonparticipants (a representative sample from the Danish population), suggesting that smokers who are motivated to quit smoking are more inclined to volunteer in a screening trial [18].

Eligible nonresponders had a higher all-cause mortality and mortality due to four other mortality classifications. However, the relative proportion of subjects that died due to all types of cancer was higher among participants. This might be explained by alcohol abuse, which is associated with higher socioeconomic status, e.g., higher educational achievement [32]. Higher alcohol consumption is associated with a higher relative risk for death from cancer [33]. Another explanation might be that participants reported more smoking-related symptoms, which may have led to more general practitioner consults. This may have led to the higher proportion of former smokers among participants and could have facilitated the detection of cancer, cardiovascular, and respiratory diseases. This may have resulted in early treatment of smoking-related diseases among NELSON participants and may have led to lower mortality rates compared with eligible nonresponders. However, the slightly younger age, better self-reported health, and healthier lifestyle among participants may have had a bigger

contribution to these differences in mortality profiles and resulted in a significantly longer follow-up among participants. As mentioned, participants were more likely to participate in any of the mentioned screening programs compared with the eligible nonresponders. Higher education levels may have led to more awareness of their risk for lung cancer and influenced the decision to participate in the NELSON trial. In addition, there were more former smokers among participants. It has been previously reported that active smoking is a barrier to participate in screening for lung cancer [20, 34]. Notable, living further from participating screening center, participants in the NELSON trial were more willing to participate than the eligible nonresponders. In contrast, the Lung-SEARCH screening trial reported that half of the responders found inability to travel the most significant reason not to participate [25].

The main strengths of this study are: (1) the large number of participants and eligible nonresponders, (2) access to all the completed first questionnaires of subjects, (3) the availability of mortality data from Statistics Netherlands, and (4) a long follow-up duration of 10 years. Finally, so far no large lung screening trial using LDCT has studied the differences in baseline characteristics and potential effect on mortality profiles between participants and eligible nonresponders. This study was limited by the fact that Statistics Netherlands could only provide aggregated mortality data. Therefore, it was not possible to perform multivariate analyses. Furthermore, all questionnaire data were self-reported, as in other studies [7]. The questionnaires included few questions on socioeconomic class and no questions on ethnic background or psychosocial profile.

In conclusion, differences in age, health, lifestyle, and socioeconomic class can lead to a healthy participant effect, i.e., a different study outcome than would have been observed if the characteristics of participants were similar to that of the target population. As expected, the distribution of participant characteristics in the NELSON study suggest that the study population is somewhat younger, healthier (e.g., more physically active, less current smokers), higher educated and more willing to participate in a screening program. These differences have influenced the mortality outcome of participants and eligible nonresponders. But, these differences are modest and therefore it seems unlikely that these differences will influence the generalizability of the main results of the NELSON trial.

ACKNOWLEDGMENTS

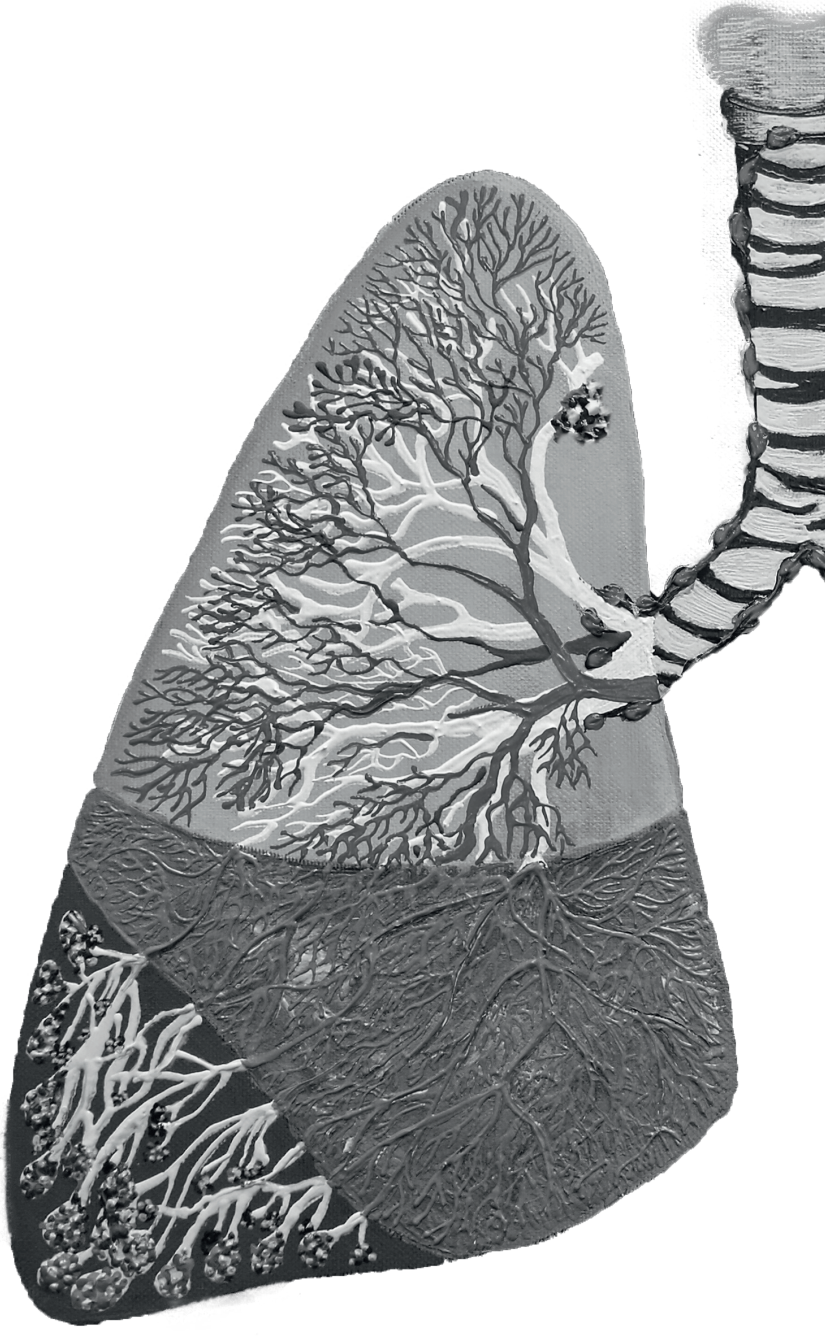
The authors thank the secretary M. Quak and the system controllers R. Faber and F.J.P. Santegoets (all from Department of Public Health, Erasmus University Medical Center) for their contribution and maintenance of the database. Further, we thank Statistics Netherlands for providing mortality data.

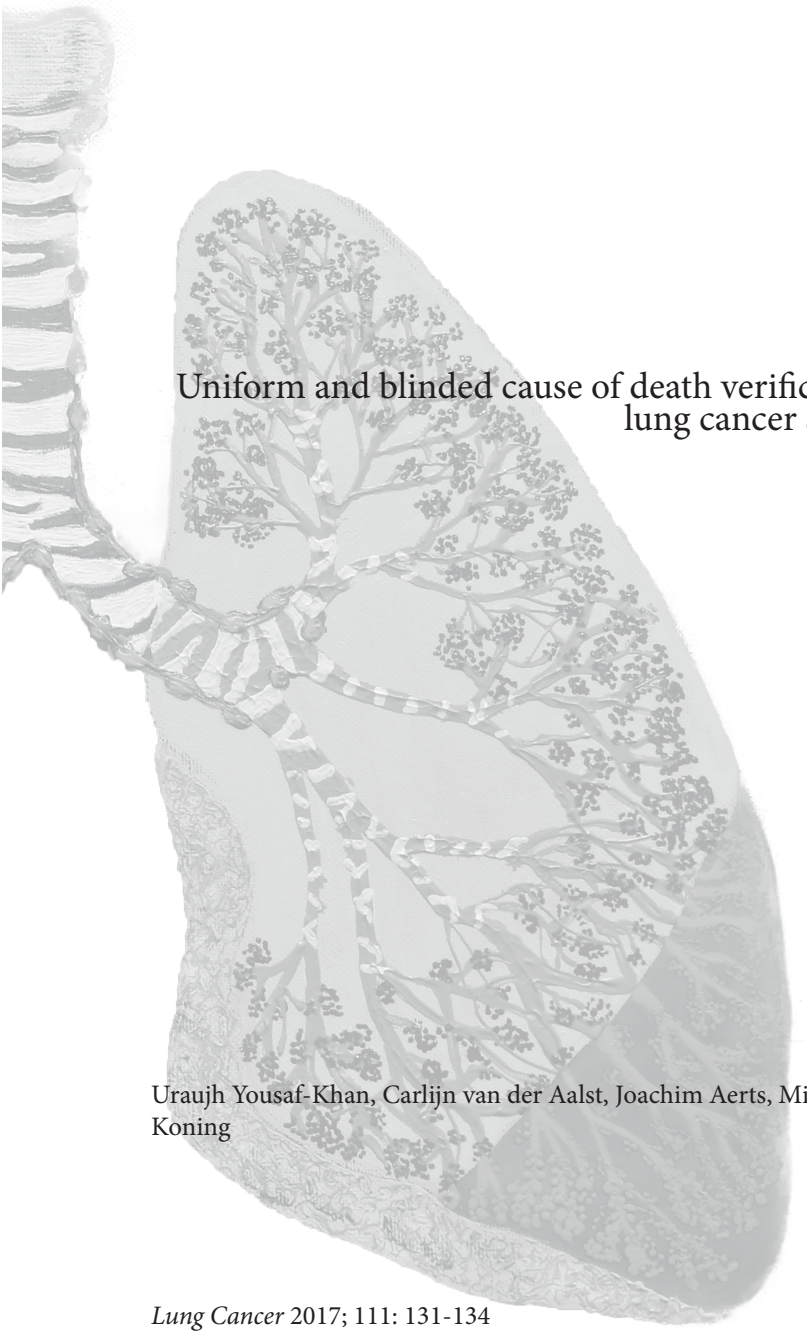
REFERENCES

1. Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. CA Cancer J Clin. 2014;64(1):9-29.
2. Gellert C, Schottker B, Brenner H. Smoking and all-cause mortality in older people: systematic review and meta-analysis. Arch Intern Med. 2012;172(11):837-44.
3. Zatonski W, Przewozniak K, Sulkowska U, West R, Wojtyla A. Tobacco smoking in countries of the European Union. Ann Agric Environ Med. 2012;19(2):181-92.
4. Tong L, Spitz MR, Fueger JJ, Amos CA. Lung carcinoma in former smokers. Cancer. 1996;78(5):1004-10.
5. Mong C, Garon EB, Fuller C, Mahtabifard A, Mirocha J, Mosenifar Z, et al. High prevalence of lung cancer in a surgical cohort of lung cancer patients a decade after smoking cessation. J Cardiothorac Surg. 2011;6:19.
6. van der Drift MA, Karim-Kos HE, Siesling S, Groen HJ, Wouters MW, Coebergh JW, et al. Progress in standard of care therapy and modest survival benefits in the treatment of non-small cell lung cancer patients in the Netherlands in the last 20 years. J Thorac Oncol. 2012;7(2):291-8.
7. National Lung Screening Trial Research T, 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.
8. Moyer VA, Force USPST. Screening for lung cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2014;160(5):330-8.
9. Infante M, Cavuto S Fau - Lutman FR, Lutman Fr Fau - Brambilla G, Brambilla G Fau - Chiesa G, Chiesa G Fau - Ceresoli G, Ceresoli G Fau - Passera E, et al. A randomized study of lung cancer screening with spiral computed tomography: three-year results from the DANTE trial. (1535-4970 (Electronic)).
10. 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.
11. Lopes Pegna A, Picozzi G. Lung cancer screening update. Curr Opin Pulm Med. 2009;15(4):327-33.
12. 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 Med. 2009;361(23):2221-9.
13. Horeweg N, Scholten ET, de Jong PA, van der Aalst CM, Weenink C, Lammers JW, et al. Detection of lung cancer through low-dose CT screening (NELSON): a prespecified analysis of screening test performance and interval cancers. Lancet Oncol. 2014;15(12):1342-50.
14. Horeweg N, van Rosmalen J, Heuvelmans MA, van der Aalst CM, Vliegenthart R, Scholten ET, et al. Lung cancer probability in patients with CT-detected pulmonary nodules: a prespecified analysis of data from the NELSON trial of low-dose CT screening. Lancet Oncol. 2014;15(12):1332-41.
15. Pinsky PF, Miller A, Kramer BS, Church T, Reding D, Prorok P, et al. Evidence of a healthy volunteer effect in the prostate, lung, colorectal, and ovarian cancer screening trial. Am J Epidemiol. 2007;165(8):874-81.

16. Dominioni L, Rotolo N, Poli A, Paolucci M, Sessa F, D'Ambrosio V, et al. Self-selection effects in smokers attending lung cancer screening: a 9.5-year population-based cohort study in Varese, Italy. J Thorac Oncol. 2010;5(4):428-35.
17. van der Aalst CM, van Iersel CA, van Klaveren RJ, Frenken FJ, 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.
18. Hestbech MS, Siersma V, Dirksen A, Pedersen JH, Brodersen J. Participation bias in a randomised trial of screening for lung cancer. Lung Cancer. 2011;73(3):325-31.
19. National Lung Screening Trial Research T, Aberle DR, Adams AM, Berg CD, Clapp JD, Clingan KL, et al. Baseline characteristics of participants in the randomized national lung screening trial. J Natl Cancer Inst. 2010;102(23):1771-9.
20. Silvestri GA, Nietert PJ, Zoller J, Carter C, Bradford D. Attitudes towards screening for lung cancer among smokers and their non-smoking counterparts. Thorax. 2007;62(2):126-30.
21. McCaffery K, Wardle J, Nadel M, Atkin W. Socioeconomic variation in participation in colorectal cancer screening. J Med Screen. 2002;9(3):104-8.
22. Miller AB, Baines CJ, To T, Wall C. Canadian National Breast Screening Study: 2. Breast cancer detection and death rates among women aged 50 to 59 years. Cmaj. 1992;147(10):1477-88.
23. Shapiro JA, Seeff LC, Nadel MR. Colorectal cancer-screening tests and associated health behaviors. Am J Prev Med. 2001;21(2):132-7.
24. Sutton S, Wardle J, Taylor T, McCaffery K, Williamson S, Edwards R, et al. Predictors of attendance in the United Kingdom flexible sigmoidoscopy screening trial. J Med Screen. 2000;7(2):99-104.
25. 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.
26. Horeweg N, van der Aalst CM, Vliegenthart R, Zhao Y, Xie X, Scholten ET, et al. Volumetric computed tomography screening for lung cancer: three rounds of the NELSON trial. Eur Respir J. 2013;42(6):1659-67.
27. Horeweg N, van Klaveren RJ, Groen HJ, Lammers JW, Weenink C, Nackaerts K, et al. Blinded and uniform cause of death verification in a lung cancer CT screening trial. Lung Cancer. 2012;77(3):522-5.
28. van Iersel CA, de Koning HJ, Draisma G, Mali WP, 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.
29. Ru Zhao Y, 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.
30. Prochaska JO, DiClemente CC. Self change processes, self efficacy and decisional balance across five stages of smoking cessation. Prog Clin Biol Res. 1984;156:131-40.
31. Davis JL, Buchanan KL, Katz RV, Green BL. Gender differences in cancer screening beliefs, behaviors, and willingness to participate: implications for health promotion. Am J Mens Health. 2012;6(3):211-7.
32. Keyes KM, Hasin DS. Socio-economic status and problem alcohol use: the positive relationship between income and the DSM-IV alcohol abuse diagnosis. Addiction. 2008;103(7):1120-30.

33. Gronbaek M, Becker U, Johansen D, Gottschau A, Schnohr P, Hein HO, et al. Type of alcohol consumed and mortality from all causes, coronary heart disease, and cancer. Ann Intern Med. 2000;133(6):411-9.
34. 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.





CHAPTER VI

Uniform and blinded cause of death verification of the NELSON lung cancer screening participants

Uraujh Yousaf-Khan, Carlijn van der Aalst, Joachim Aerts, Michael den Bakker, Harry de Koning

Lung Cancer 2017; 111: 131-134

ABSTRACT

Background
Primary outcome of the Dutch-Belgian lung cancer screening trial (NELSON) is lung cancer-specific mortality. Accurate assessment of the cause of death (CoD) is crucial. As death certificates regarding the CoD can be inaccurate, a clinical expert committee (CEC) was formed to assign the CoD. In this study, the medical files of deceased lung cancer patients were reviewed and the outcomes were compared with official death certificates.

Methods
The first 266 completed medical files of Dutch deceased participants who were diagnosed with lung cancer during the study or of those with lung cancer on the death certificate were selected and blinded towards arms and patients identity. The end product of the review process consisted of six possible categories which defined the graduation of certainty that lung cancer was the primary CoD. The percentage agreement and the Cohen's kappa statistics between the two CEC-members were calculated. The sensitivity and specificity of the official death certificates were determined.

Results
The results indicated that, the overall concordance and the Cohen's kappa between the CEC-members were 86.1% and 0.57 (0.45–0.69, $p < 0.001$), respectively. This level increased with the numbers of cases evaluated. The sensitivity and the specificity of the official death certificate were 92.6% and 98.8%; 6.5% cases were reclassified to lung cancer specific death, which is lower than in the National Lung Screening trial (22.0%).

Conclusions
Concluding, each death should be reviewed by at least two members. So far, in the NELSON trial, possible biases related to lung cancer death seem relatively small.

6.1. INTRODUCTION

In cancer screening trials, the cause of death (CoD) is often determined by a review committee of medical scientists who (independently) review the blinded medical files of the deceased participants and achieve consensus regarding the underlying cause of death [1-4]. This is done to overcome different biases, as sticky-diagnosis (if more cancers are being diagnosed in the screening group, it's likely that more death are attributed to that cancers compared to the control group with no screening), and slippery-linkage bias (where deaths are due to the screening process, but are not traced back to screening and are certified as due to other causes). This also should overcome the variable sensitivity and specificity of the official death certificates which depends on the accuracy of the certifying clinicians [5-7].

A clinical expert committee (CEC) was formed to independently, and in a uniform and blinded matter to review the first 266 completed medical files of deceased lung cancer participants of the NELSON trial – the Dutch Belgian lung cancer screening trial [8, 9]. Furthermore, these files and official death certificates were compared.

6.2. METHODS

Pilot study

Previously, a pilot study (n=50) by a uniform classification (review committee vs. death certificates) demonstrated an agreement of 90% (Cohen's kappa 0.65) [9]. The sensitivity and specificity of the official death certificates for lung cancer specific mortality were 95.2% and 62.5%, respectively, what implied that the final NELSON outcomes should be established with predetermined criteria and an independent review of blinded cases.

Selection of subjects

For this study, all Dutch deceased participants who were diagnosed with lung cancer (during the study or at autopsy), deceased participants who were in the diagnostic work-up for lung cancer, and participants with a notation of lung cancer on the death certificate (International Classification of Diseases and related health problems (ICD) version 10:34) were selected. Lung cancer cases and the death certificates were obtained through linkages with the National Cancer Registry of the Netherlands (100% coverage), and Statistics Netherlands (100% coverage; 2003–2014), respectively. Thereafter, it was verified if the participant had indeed been diagnosed with lung cancer, during a separate procedure. All relevant medical data pertaining to the CoD (all outpatient and discharge letters, radiology and pathology reports, and place of death) was collected.

The NELSON clinical expert committee

The CEC consisted of an independent pulmonologist-oncologist and a pathologist specialized in lung oncology. In cases with no consensus, a third reviewer (a clinical epidemiologist specialized in screening) was consulted[9].

The CoD review process

After blinding the medical files for patients identity and study arm, they were uploaded onto a secure online database. The end product of the evaluation consisted of six possible categories which defined the graduation of certainty that lung cancer was the primary cause of death, which is based upon the CoD review process followed in the European Randomized Screening for Prostate Cancer trial [10]. After reaching consensus between the two CEC-members, and the third reviewer, the end product was considered as the golden standard (**Figure 1**). At all time, the reviewers had no access to the official death certificates.

After the evaluation, the nonconsensual cases were discussed plenary to reach consensus. If it was not possible to reach consensus a third reviewer was consulted. After reaching consensus, first, it was checked if the files were also reviewed in the pilot study [9]. If yes, the consensus of both studies were compared. The nonconsensual cases, were re-evaluated by the committee and the clinical epidemiologist from the pilot study. The final agreement after the meeting was considered as the golden standard. In case the files were not reviewed in the pilot study, the consensus reached by the committee was considered the golden standard. After reaching consensus over all cases, the final outcome was compared with the primary CoD on the official death certificate by the first author. In case the CoD on the official death certificate was 'lung cancer death' and the CoD reviewed CoD was 'death due other cause of death than lung cancer', the committee and the clinical epidemiologist reviewed these files again and discussed it in a meeting.

Analysis

The primary CoD was defined as 'the disease that initiated the chain of morbid events directly leading to death'. Lung cancer mortality, was defined as 'definitely' or 'probable lung cancer death'. All other four possible categories ('possible', 'unlikely', 'definitely no lung cancer death', and 'intercurrent death with lung cancer as contributing factor') were considered as 'another CoD'. CEC-members Cohen's Kappa represented the percentage agreement between the CEC-members. Sensitivity (true positives (lung cancer death assigned by both sources) divided by the sum of true positives and false negative diagnoses according to the official death certificates) and specificity (true negatives (other cause of death assigned by both sources) divided by the total death due to other cause according to the official death certificates) were calculated. All continuous variables were presented as medians and interquartile ranges (IQR), as appropriate. Differences between variables were calculated by using the Median Test or ANOVA (continuous variables), chi-square test (nominal), and Kruskall-Wallis test (categorical). For all analysis SPSS version 21 was used.

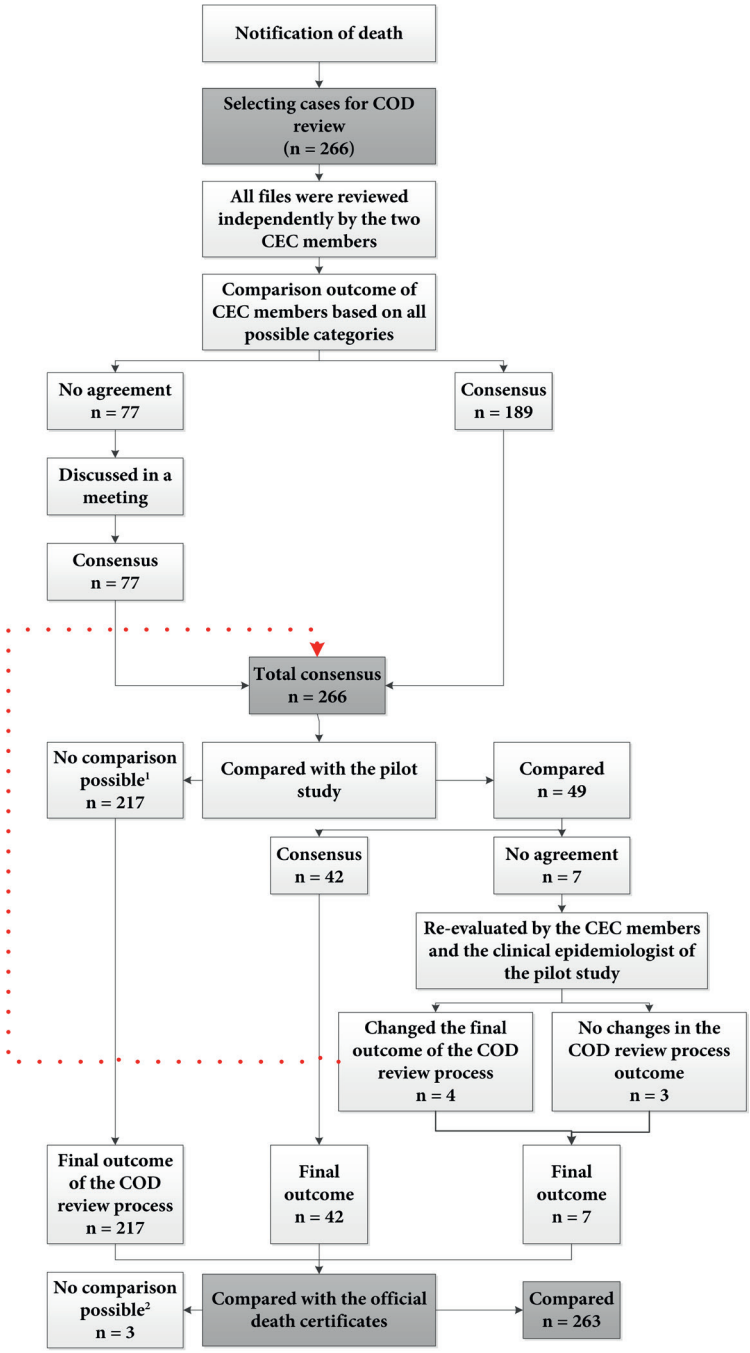


Figure 1. an overview of the CoD review process.

[1]These 217 cases were not reviewed in the pilot study and therefore no comparison was possible with the pilot study.
[2] In these 3 cases the participants did not give permission to obtain their official death certificate. Therefore no comparison with the official death certificate was possible.
Red dotted line: after discussing 7 cases in a meeting with a third reviewer from the pilot study and the two CEC members, in 4 cases the initial outcome of the CoD review process was changed.

6.3. RESULTS

In total, 266 medical files of Dutch deceased participants, who deceased between 28th August 2004 and 26th April 2014, were selected for the review (**Table 1**).

Agreement between the CEC-members based on six possible categories was reached in 71.1% (189/266) of the online reviewed cases. Divided in a lung cancer death (definitely or probably lung cancer death), agreement between the two members was reached in 86.1% of the cases (Cohen's kappa of 0.57 (95% CI, 0.45–0.69, $p < 0.001$)). Nonconsensual cases were discussed in a meeting. Reasons of disagreement were: recently received lung cancer treatment (n=1), autopsy revealed lung cancer death (n=2), other possible CoD (n=2), too little information (n=2), intercurrent death with another CoD (n=2), another CoD was more obvious (n=2), and reports were available about progressive tumor (n=26). No third reviewer was required to reach consensus.

Compared to the pilot study (in which 49 comparable cases were evaluated by two different reviewers), in seven out of 49 cases there was a disagreement (Fig. 1). These cases were re-evaluated and discussed in a meeting by the CEC and the third reviewer. In four cases the outcome of the CoD review process was changed to 'definitely lung cancer death'. Reasons for these changes were: autopsy showed lung cancer death (n=2), euthanasia because of cerebral metastasis of lung cancer (n=1) and all the clinicians who treated the patient addressed the death of the patient to lung cancer (n=1).

From 266 participants, three participants did not provide informed consent to obtain their official death certificate. Therefore, the consensus reached in the CoD review process will be used as the primary CoD for these cases.

When the official death certificates noted 'lung cancer death', but the CoD review process noted it as another CoD', a third independent review took place (n=21). In three cases the outcome of the COD review committee was changed to 'lung cancer death'. In all the three cases the patient suffered from metastasized lung cancer with no treatment options. The overall sensitivity and specificity of the death certificates were 92.6% and 98.8%, respectively. Death review resulted in a reclassification of 12.2% (32/263) of the cases (**Table 1**).

Table 1: Causes of death by the reviewers after consensus meeting and the official certificates

Death certificates	Cause of death Review committee		Total
	LC death	Other cause of death	
	n (%)	n (%)	
LC death	212 (80.6)	17 (6.5)	229 (87.1)
Other cause of death	15 (5.7)	19 (7.2)	34 (12.9)
Total	227 (86.3)	36 (13.7)	263 (100.0)

LC: lung cancer
LC death: definitely LC death or probable LC death
Other cause of death: possible LC death, unlikely LC death, definitely no LC death and intercurrent death with LC as contributing factor

Reasons for reclassification of "other CoD" to "lung cancer death" were: cases with severe metastatic lung cancer (n=8), lung cancer diagnosis shortly before death (n=5), died during work-up for lung cancer and autopsy showed lung cancer death (n=1) and died from complication after a thoracotomy for lung cancer treatment (n=1).

Reasons for reclassification of "lung cancer death" to "other CoD" were: other malignancy (n=7), multiple cardiovascular and respiratory problems (n=3), cardiovascular problems (n=3), too little information (n=1) and other causes (n=3).

6.4. DISCUSSION

In this study, medical files of deceased NELSON lung cancer participants were reviewed concerning the underlying CoD and were compared with the official death certificates. Weak to moderate agreement of 86.1% (Cohen's kappa of 0.57) between the two committee members was reached, what underlies the need for a CoD committee. As expected, the level of agreement increased with the numbers of cases evaluated (data not shown) [9].

The National Lung Screening Trial (NLST) [3, 4], which also used a death review committee to verify the CoD, showed a comparable sensitivity and specificity of the official death certificates (NLST: 91% and 97%, and NELSON 92.7% and 98.8%, respectively). However, the NLST review was not restricted to the files of lung cancer patients only [4]. Secondly, only one member reviewed the medical files of the NLST first. In case of concordance with the death certificate this was considered as certified. In the NELSON trial, two members reviewed each file, independently. Furthermore, less official death certificates were re-classified in the NELSON trial compared to the NLST (12.2% vs. 26.0%). The NLST showed a higher reclassification of deaths to lung cancer specific deaths than the NELSON trial (22.0% vs. 6.5%). Dutch clinicians and Statistics Netherlands may have a more uniform method in reporting the cause of death.

Furthermore, the NLST showed that the mortality benefit of screening did not significantly change with reviewing the CoD [4]. For the NELSON trial, we are waiting the final results and the CoD review process will be continued until then.

The National Lung Screening Trial (NLST) [3, 4], which also used a death review committee to verify the CoD, showed a comparable sensitivity and specificity of the official death certificates (NLST: 91% and 97%, and NELSON 92.7% and 98.8%, respectively). However, the NLST review was not restricted to the files of lung cancer patients only [4]. Secondly, only one member reviewed the medical files of the NLST first. In case of concordance with the death certificate this was considered as certified. In the NELSON trial, two members reviewed each file, independently. Furthermore, less official death certificates were re-classified in the NELSON trial compared to the NLST (12.2% vs. 26.0%). The NLST showed a higher reclassification of deaths to lung cancer specific deaths than the NELSON trial (22.0% vs. 6.5%). Dutch clinicians and Statistics Netherlands may have a more uniform method in reporting the cause of death. Furthermore, the NLST showed that the mortality benefit of screening did not significantly change with reviewing the CoD [4]. For the NELSON trial, we are waiting the final results and the CoD review process will be continued until then.

Strengths of this study are the blinded (for study arm and official death certificate) and uniform review process by the independent reviewers that had access to participant's complete medical file. A potential limitation is the selection of the first 266 deceased participants from whom all medical data was collected, what possibly over represents participants with an aggressive lung cancer leading to more assignable death due to lung cancer and few cases to be reclassified. Secondly, a subset of deaths was selected for the CoD review (e.g. participants with a diagnosis of lung cancer or with a notation of lung cancer on the death certificate), what may increase the likelihood of lung cancer as cause of death. Furthermore, participants with another CoD were excluded, what introduces some uncertainty about the generalizability of the results for the whole deceased NELSON population. However, reviewing all deaths is too time-consuming.
In conclusion, for large randomized cancer screening trials it is necessary to review the cancer-specific cause of death because the official death certificate report's varying accuracy (sensitivity and specificity). In the NELSON trial, the results indicated that possible biases related to lung cancer death seem relatively small. Furthermore, it is recommended that a uniform and blinded death review should be done by at least two independent members.

ACKNOWLEDGMENTS

The authors thank the system controllers R. Faber and F.J.P. Santegoets, and the secretary M.Quak (all from the department of Public Health, Erasmus University Medical Center Rotterdam for their contribution and maintenance of the database. Furthermore, we thank R. Ziengs (University Medical Center Groningen), S. van Amelsvoort-van der Vorst (University Medical Center Utrecht) and M.S.G. den Uijl (Erasmus Medical University Center). Finally, we thank the Dutch Cancer Registry (NKR) and the Statistics Netherlands (CBS) for the data linkages.

REFERENCES

1. Williams NJ, Hill EM, Ng SY, Martin RM, Metcalfe C, Donovan JL, et al. Standardisation of information submitted to an endpoint committee for cause of death assignment in a cancer screening trial - lessons learnt from CAP (Cluster randomised triAl of PSA testing for Prostate cancer). BMC Med Res Methodol. 2015;15:6.
2. Otto SJ, van Leeuwen PJ, Hoekstra JW, Merckelbach JW, Blom JH, Schroder FH, et al. Blinded and uniform causes of death verification in cancer screening: a major influence on the outcome of a prostate cancer screening trial? Eur J Cancer. 2010;46(17):3061-7.
3. Marcus PM, Gareen IF, Miller AB, Rosenbaum J, Keating K, Aberle DR, et al. The National Lung Screening Trial's Endpoint Verification Process: determining the cause of death. Contemp Clin Trials. 2011;32(6):834-40.
4. Marcus PM, Doria-Rose VP, Gareen IF, Brewer B, Clingan K, Keating K, et al. Did death certificates and a death review process agree on lung cancer cause of death in the National Lung Screening Trial? Clin Trials. 2016;13(4):434-8.
5. Maudsley G, Williams EM. Death certification by house officers and general practitioners--practice and performance. J Public Health Med. 1993;15(2):192-201.
6. Degani AT, Patel Rm Fau - Smith BE, Smith Be Fau - Grimsley E, Grimsley E. The effect of student training on accuracy of completion of death certificates. (1087-2981 (Electronic)).
7. Aung E, Rao C Fau - Walker S, Walker S. Teaching cause-of-death certification: lessons from international experience. (1469-0756 (Electronic)).
8. 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 Med. 2009;361(23):2221-9.
9. Horeweg N, van Klaveren RJ, Groen HJ, Lammers JW, Weenink C, Nackaerts K, et al. Blinded and uniform cause of death verification in a lung cancer CT screening trial. Lung Cancer. 2012;77(3):522-5.
10. De Koning HJ, Blom J Fau - Merkelbach JW, Merkelbach Jw Fau - Raaijmakers R, Raaijmakers R Fau - Verhaegen H, Verhaegen H Fau - Van Vliet P, Van Vliet P Fau - Nelen V, et al. Determining the cause of death in randomized screening trial(s) for prostate cancer. (1464-4096 (Print)).

APPENDIX I: FULL ARTICLE

Uniform and blinded cause of death verification of the NELSON lung cancer screening participants

Uraujh Yousaf-Khan, Carlijn van der Aalst, Joachim Aerts, Michael den Bakker, Harry de Koning

ABSTRACT

Background
The primary outcome of the Dutch-Belgian lung cancer screening trial (NELSON) is lung cancer-specific mortality for which an accurate assessment of the cause of death (CoD) is crucial. Because death certificates regarding the cause of death can be inaccurate, a clinical expert committee (CEC) was formed to assign the CoD. In this study, the medical files of deceased lung cancer patients were reviewed and the outcomes were compared with official death certificates.

Methods
The first 266 completed medical files of Dutch deceased participants who were diagnosed with lung cancer during the course of study or of those with lung cancer on the death certificate were selected and blinded towards arms and patients identity. The end product of the review process consisted of six possible categories which defined the graduation of certainty that lung cancer was the primary CoD. The percentage agreement and the Cohen's kappa statistics between the two CEC members were calculated. The sensitivity and specificity of the official death certificates were determined.

Results
The overall concordance and the Cohen's kappa between the CEC members were 86.1% and 0.57 (0.45-0.69, $p<0.001$), respectively. Level of agreement between the two CEC members increased with the numbers of cases they evaluated. The sensitivity and the specificity of the official death certificate were 92.6% and 98.8%; 6.5% cases were reclassified to lung cancer specific death, which is lower than in the National Lung Screening trial (22.0%).

Conclusion
It is recommended that death review of each case should be determined by at least two members, independently and in a uniform and blinded matter. So far, in the NELSON study, possible biases related to lung cancer death seem relatively small.

INTRODUCTION

Lung cancer mortality is the primary outcome of lung cancer screening trials that investigate the (cost)-effectiveness of a cancer screening programme [1-3]. In cancer screening programmes, the cause of death (CoD) is often determined by a review committee [4-6]. The review committee usually consists of clinicians or other medical scientists who (independently) review the blinded medical files of the deceased participants and achieve consensus regarding the underlying cause of death (UCoD). This is done to overcome different biases, as sticky-diagnosis (where more cancers are being diagnosed in the screening group and therefore deaths are more likely to be attributed to that cancer compared to the control group with no screening), and slippery-linkage bias (where deaths are due to the screening process, e.g. a diagnostic thoracotomy, which are not traced back to screening but are certified as due to other causes and may lead to an overestimation of the beneficial effects of screening). Also, to overcome the variable sensitivity and specificity of the official death certificates which depends on the accuracy of the certifying clinicians [7-9].

The largest lung cancer screening trial in Europe is the Dutch-Belgian lung cancer screening trial (NELSON), which aims to investigate whether LDCT screening reduces lung cancer mortality by 25% or more compared to no screening [3, 10, 11]. A clinical expert committee (CEC) was formed to review the medical files of the deceased NELSON lung cancer participants, independently, and in an uniform and blinded fashion [12].

The aim of this study is to evaluate the first 268 reviewed medical files of the deceased NELSON lung cancer participants by the CEC, as well as the comparison between these files and official death certificates.

METHODS

The NELSON trial

In brief, the NELSON trial is a sufficiently powered, randomized-controlled lung cancer screening trial with the aim to investigate whether LDCT screening reduces lung cancer mortality by ≥25% in high-risk subjects for developing lung cancer compared to no screening [3, 13, 14]. Initially, 15,822 participants were randomized (1:1). However, through linkages with the national cancer registries of the Netherlands and Belgium, and the death registries of the Netherlands and Belgium (Statistics Netherlands and the Flemish Agency for Care and Health, respectively), it appeared that 30 participants died before randomization and should therefore be ruled out from further analysis; 15 of these belonged to the screening group and 15 to the control group. The actual number of randomization are: 15,792 participants (7,900 in to a screening group and 7,892 in to a control group). LDCT screening took place at baseline, and after one, two and two-and-a-half years. Control group received usual care.

In our previous report [12] , a CoD review process protocol was developed and tested in a pilot study. In brief, in the pilot study the medical files of the first 50 deceased NELSON lung cancer participants were reviewed independently by two members. The files were blinded for patients identity and study arm. The final outcome after the review process was compared with the CoD according to the official death certificates. An agreement of 90% (Cohen's kappa 0.65) was seen, which demonstrated a uniform classification. The sensitivity and specificity of the death certificates for lung cancer specific mortality were 95.2% and 62.5%, respectively. Furthermore, it implied that the final outcome of the NELSON trial should be established with predetermined criteria and an independent review of blinded cases.

Identification of the deceased participants

Subjects of the CoD review process were: 1) all deceased participants who were diagnosed with lung cancer (during the study or at autopsy), 2) deceased participants who were in the diagnostic work-up for lung cancer, and 3) participants with a notation of lung cancer on the death certificate (International Classification of Diseases and related health problems (ICD) version 10:34). The lung cancer cases were identified through linkages with the national cancer registry of the Netherlands (100% coverage), and the death certificates were obtained (100% coverage) from the Statistics Netherlands from 2003 until 2014. For all identified subjects, firstly it was verified whether the participant had indeed been diagnosed with lung cancer during the course of study (e.g. after randomization) or at autopsy. This verification process was performed separately from the CoD review process and will not be addressed in this manuscript. Hereafter, all relevant medical data pertaining to the CoD was collected. The relevant medical data consisted of: all information provided by the general practitioner, outpatient visit letters, discharge letters, reports of radiology, nuclear medicine, pathology and microbiology, laboratory results, autopsy reports, and date and place of death.

The clinical expert committee of the NELSON trial

The CEC consisted of a pulmonologist-oncologist and a pathologist specialized in lung oncology (both were not involved in the NELSON trial or in care of the patients). In cases were no consensus was reached, a third reviewer (a clinical epidemiologist specialized in screening) was consulted [12].

The CoD review process

The medical files were blinded for participant's identity and study arm by an individual who was not a member of the CEC. Hereafter, the files were uploaded onto a secured online database so that each member of the CEC could independently review the files. The evaluation process performed by the members was guided by the use of a decisional flowchart and a detailed list of criteria that guide the decision-making process uniformly, which has been used in the pilot study (both available in the appendix) [12]. The end product of the evaluation consisted of six possible categories which defined the graduation of certainty that lung cancer was the

primary cause of death (supplementary material, **Table 1**). This graduation of certainty is based upon the CoD review process followed in the European Randomized Screening for Prostate Cancer trial (ERSPC) [15].

After completing the evaluation online, the nonconsensual cases were discussed plenary to reach consensus. Next, it was first checked if the files were also reviewed in the pilot study [12]. If this was the case, the consensus of the CoD review process was compared with the consensus reached in the pilot study. In case of no agreement, the files were re-evaluated by the two CEC members and by the clinical epidemiologist from the pilot study and discussed in a meeting. The agreement after the meeting was considered as the golden standard. In case the files were not reviewed in the pilot study, the consensus reached by the two CEC members was considered the golden standard.

After reaching consensus over all cases, the final outcome of the CoD review process was compared with the primary CoD on the official death certificate by the first author. In case the CoD on the official death certificate was 'lung cancer death' and the CoD reviewed CoD was 'death due other cause of death than lung cancer', the two CEC members and the clinical epidemiologist reviewed these files again and discussed it in a meeting. At all time, the reviewers had no access to the official death certificates.

Analysis

For this study, the first 268 completed medical files of Dutch deceased participants with lung cancer were included. The primary cause of death was defined as 'the disease that initiated the chain of morbid events directly leading to death'. The primary endpoint of the study, lung cancer mortality, was defined as 'definitely' or 'probable lung cancer death'. All other four possible categories ('possible', 'unlikely', 'definitely no lung cancer death', and 'intercurrent death with lung cancer as contributing factor') were considered as 'death due to other causes'. Agreement between the two CEC members was expressed in percentage agreement in Cohen's Kappa [16]. A Cohen's kappa of 1 and 0 indicated a 'perfect agreement' and 'no agreement', respectively. Sensitivity and specificity of the death certificates compared with the CoD review process consensus were calculated. Sensitivity: proportion of true positives (lung cancer death assigned by both sources) divided by the sum of true positive and false negative diagnoses according to the official death certificates. Specificity: proportion of true negatives (other cause of death assigned by both sources) divided by the total death due to other causes according to the official death certificates. None of the continuous variables tested by using the Kolmogorov-Smirnov test and examining Q-Q-plots were distributed normally. Therefore, the variables were described by using medians and interquartile ranges (IQR). Differences between the continuous variables across the groups were calculated by using the Median Test or ANOVA. Differences between nominal variables were calculated by using a chi-square test. Differences between categorical variables were tested by using Kruskall-Wallis Test. For all analysis IBM SPSS version 21 was used.

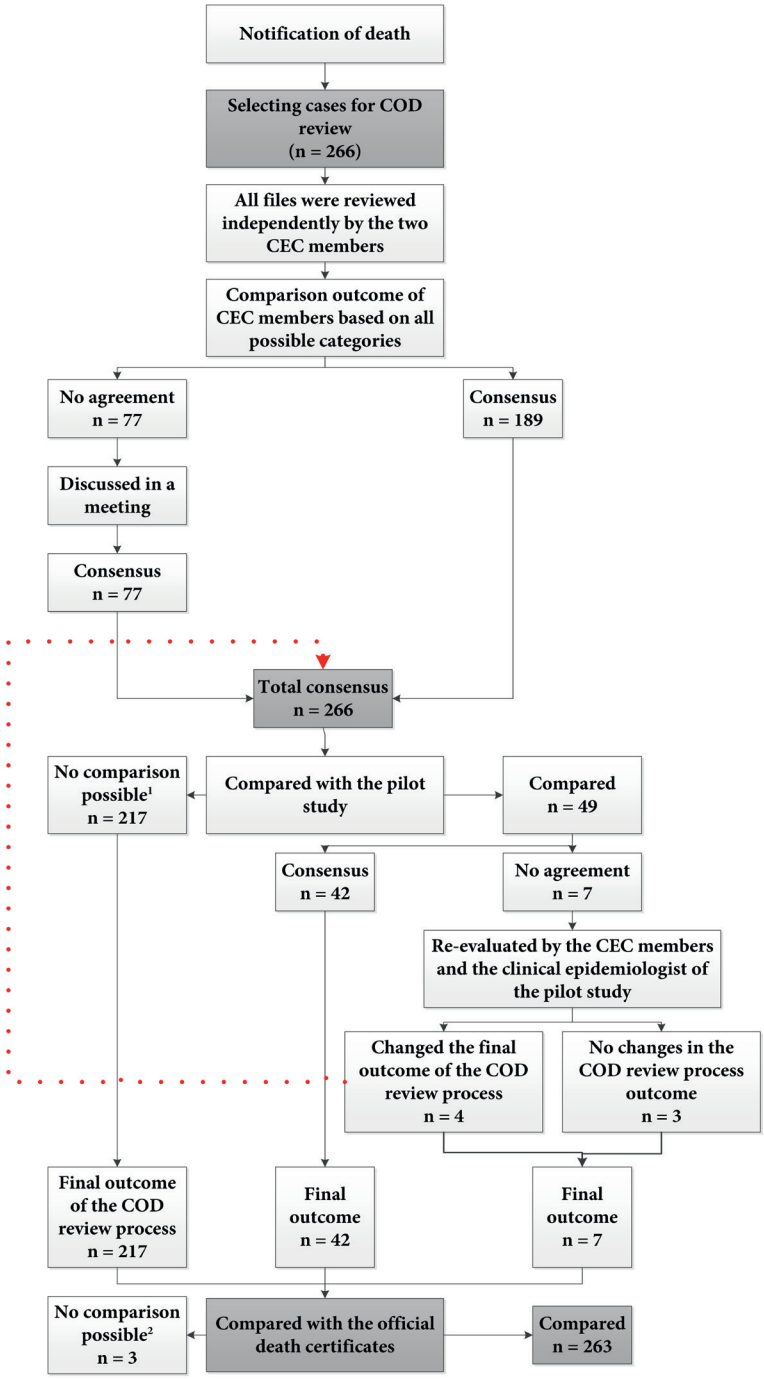


Figure 1: an overview of the CoD review process

[1]These 217 cases were not reviewed in the pilot study and therefore no comparison was possible with the pilot study.

[2] In these 3 cases the participants did not give permission to obtain their official death certificate. Therefore no comparison with the official death certificate was possible.

Red dotted line: after discussing 7 cases in a meeting with a third reviewer from the pilot study and the two CEC members, in 4 cases the initial outcome of the CoD review process was changed.

RESULTS

In total, 266 medical files of Dutch deceased participants were selected and reviewed by the CEC. The selected medical files showed that the participants died between 25th April 2004 and 26th April 2014. According to the information received from Statistics Netherlands only, in the same period 1,589 NELSON participants (10%) deceased (all-cause mortality).

The selected deceased participants were slightly older compared to the non-deceased participants at baseline (62.1 years vs. 58.0 years), less often females (12.8% vs. 17.3%), more often current smokers at baseline (67.3% vs. 54.9%) and had smoked more pack-years at baseline (43.7 years vs. 38.0 year; **Table 1**).

For 248 out of 266 cases, the place of death was known. In 65.3% (166/248) the place of death was outside the hospital: 75.3% (125/166) at home, 13.3% (22/166) in a hospice and 9.0% (15/166) in a nursing home. In 83.1% (221/266) of the cases it was known whether an autopsy was performed: 4.5% (10/221) had no autopsy compared to 95.5% (211/221) with autopsy. In 45.1% (120/266) of the cases it was known whether the patient received palliative sedation or euthanasia; in 12.5%(15/120) palliative sedation was given and in 8.3% (10/120) euthanasia was performed.

Table 1. characteristics of non-deceased NELSON participants vs. deceased NELSON participants

	Non-deceased participants[1]	Selected deceased participants	Other deceased participants	p-value
N	14,200	266	1,326	
Age at baseline, years				<0.001
Median	58.0	62.1	62.1	
IQR	7.8	9.1	9.5	
Age at death, years				0.08
Median	NA	67.0	67.6	
IQR	NA	9.5	10.2	
Gender, n (%)				<0.001
Female	2,454 (17.3)	34 (12.8)	105 (7.9)	
Smoking status at baseline, n (%)				
Smokers	7,793 (54.9)	179 (67.3)	776 (58.5)	<0.001
Pack-years smoked at baseline, years (IQR)	38.0 (19.8)	43.7 (26.0)	39.2 (24.0)	0.002

[1]For these analysis non-deceased NELSON participants were defined as 15,792 randomized participants – all deceased participants.

Agreement between the reviewers of the CEC

In **Table 2a** an overview is presented of all the possible outcomes of separate review of the CoD of each CEC member before a consensus meeting took place. In 71.1%(189/266) of the cases there was an agreement between the two CEC members. The cases with disagreement (n=77) were discussed in a meeting. No third review was required to reach consensus in any of the cases.

Table 2b, presents an overview of separate review of the CoD by each CEC member, before a consensus meeting took place, divided in a lung cancer specific death (definitely or probably lung cancer death) and in other CoD (possible, unlikely or definitely no lung cancer death, and intercurrent death with lung cancer as contributing factor). Agreement between the two CEC members was reached in 86.1% of the cases, which corresponds with a Cohen's kappa of 0.57 (95%CI, 0.45-0.69, p<0.001). In 37 (28+9) cases no agreement was reached and these cases were discussed by the two CEC members in a consensus meeting. Reasons of disagreement are pointed out in **Table 3**.

Table 2a: outcome of separate review of the cause of death, before consensus meeting (all categories)

		Reviewer 2						
Reviewer 1		Definitely LC death	Probable LC death	Possible LC death	Unlikely LC death	Definitely no LC death	Intercurrent death with LC as contributing factor	Total
	Definitely LC death	147	10	5	1	0	3	166 (62.4)
	Probable	23	14	0	0	0	0	37 (13.9)
	Possible	13	3	7	1	1	2	27 (10.2)
	Unlikely	1	0	0	1	0	0	2 (0.8)
	Definitely no LC death	3	1	1	0	15	2	22 (8.3)
	Intercurrent death with LC as a contributing factor	7	0	0	0	0	5	12 (4.5)
	Total	194 (72.9)	28 (10.5)	13 (4.9)	3 (1.1)	16 (6.0)	12 (4.5)	266 (100.0)

Table 2b: outcome of separate review of the cause of death, before consensus meeting (two possible categories[1])

	Reviewer 2		
Reviewer 1	LC death	Other cause of death	Total
LC death	194 (72.9)	9 (3.4)	203 (76.3)
Other cause of death	28 (10.5)	35(13.2)	63 (23.7)
Total	222 (83.5)	44 (16.5)	266 (100.0)

LC: lung cancer

[1]Two possible categories were:

- LC death: definitely LC death or probable LC death
- Other cause of death: possible LC death, unlikely LC death, definitely no LC death and intercurrent death with LC as contributing factor.

Cohen's Kappa: 0.57 (95%CI, 0.45-0.69, p<0.001)

Table 3: reasons of disagreement between the two CEC members

Reasons	N	%
Reports about progressive tumor or stage IV tumor were available until death	26	70.3
Underwent recently lung cancer treatment	1	2.7
Autopsy reports showed lung cancer as CoD	2	5.4
Other CoD are possible according to the medical data	2	5.4
Too little information	2	5.4
Intercurrent lung cancer death with another CoD	2	5.4
Another CoD was clear	2	5.4
Total	37	100

CoD: cause of death

After reaching consensus, the files were compared with reviewed files from the pilot study. In seven out of 49 cases there was no agreement between the outcome of the CoD review process and the pilot study (**Figure 1**). Therefore, these seven cases were re-evaluated and discussed in a meeting by the two CEC members and by the clinical epidemiologist from the pilot study. In four cases the outcome of the CoD review process was changed from 'intercurrent death with lung cancer as contributing factor (2x)', 'other cause of death' and 'possible LC death' to 'definitely lung cancer death'. Reasons for these changes were: in two cases an autopsy was performed with lung cancer death as the CoD, in one case euthanasia was performed because of cerebral metastases of lung cancer and in one case all the specialists who treated the patients addressed the death of the patients from lung cancer.

Comparison of the reviewed medical files with the official death certificates

From 266 participants, three participants did not provide informed consent to obtain their official death certificate. Therefore, no comparison of the medical files of these 3 participants could be made with the official death certificates. For these three cases, the consensus reached in the CoD review process will be used as the primary CoD.

In cases in which the official death certificates noted 'lung cancer death', but the CoD review process noted it as 'death due to other cause', a third independent review took place (n= 21). In four cases the outcome of the COD review committee was changed to 'lung cancer death'. Eventually, in 12.2% (32/263) of the cases there was no concordance with the official death certificate.

The overall sensitivity and specificity of the death certificates were 92.6% and 98.8%, respectively. Death review resulted in a reclassification of 12.2% of the cases (**Table 4**). In

15 out of these 32 (46.9%) cases, the CEC reclassified the CoD to a lung cancer specific death instead of 'other CoD' as noted on the official death certificate. Reasons of this reclassification: eight participants had a severe metastatic lung cancer, five participants were diagnosed with lung cancer shortly before death, in one case the participant was in work-up for lung cancer and the autopsy showed the cause of death was lung cancer death and one participant died from complications after a thoracotomy for lung cancer treatment. In the other 17 cases (53.5%), the CEC reviewed the CoD as 'another CoD', while according to the official death certificate the UCD was lung cancer death. The other CoDs were: another malignancy was also present (7x), cardiac cause of death (3x), multiple (lung) problems (3x), too little information of the final phase (1x) and other clear causes (3x).

Table 4: causes of death by the reviewers after consensus meeting and the official certificates

Death certificates	**Cause of death Review committee**		**Total**
	LC death	Other cause of death	
	n (%)	n (%)	
LC death	212 (80.6)	17 (6.5)	229 (87.1)
Other cause of death	15 (5.7)	19 (7.2)	34 (12.9)
Total	227 (86.3)	36 (13.7)	263 (100.0)

LC: lung cancer
LC death: definitely LC death or probable LC death
Other cause of death: possible LC death, unlikely LC death, definitely no LC death and intercurrent death with LC as contributing factor

DISCUSSION

In this study, medical files of deceased NELSON lung cancer participants were reviewed and compared with the official death certificates. An agreement between the two CEC members of 86.1% (Cohen's kappa of 0.57) was reached. The outcome of the review process was compared with the official death certificates: the overall sensitivity and the specificity of the official death certificates were 92.6% and 98.8%, respectively.

This study showed a weak to moderate agreement between the two CEC members before a consensus meeting took place. This addresses the need of more than one reviewer, and should be recommended for end-point verifications in large screening trials with cancer-specific mortality as outcome. In cases with disagreement, all relevant medical data was available to reach consensus. In accordance with our expectations, the level of agreement slightly increased with the numbers of cases they evaluated: the so-called 'learning effect' (data not shown) [12].

The National Lung Screening Trial (NLST), also used a death review committee to verify the UCoD of the deceased participants [6, 17]. The NLST reviewed 42% of the deaths (1,642 out of

3,877) and showed a sensitivity of 91% and a specificity of 97% of the official death certificates. This is comparable with the NELSON trial. However, their endpoint verification differed from the NELSON trial. The NLST used a broader selection of participants whose medical files were reviewed: not only participants with a notation of lung cancer on the death certificate and of those occurring among participants ever diagnosed with lung cancer were selected, but, for example, also participants with a cancer of interest, that might be misdiagnosed as no lung cancer [17]. Secondly, after selecting the cases of interest, the first step in the NLST was that only one member reviewed the medical files. If there was concordance with the death certificate the reviewed CoD was considered as certified. However, in the NELSON trial each medical file was reviewed independently by two members of the CEC. Also, so far, in the NELSON trial 12.2% of official death certificates were re-classified, while in the NLST it was 26.0%. The NLST showed a higher reclassification of deaths to lung cancer specific deaths than the NELSON trial (22.0% vs. 6.5%). This may suggest that in the Netherlands a higher accuracy of clinicians is seen regarding reporting deaths related to lung cancer or due to other causes of death. However, in the NELSON trial, so far, fewer cases have been reviewed. Furthermore, the NLST showed that the mortality benefit of screening did not significantly change with reviewing the CoD. As the primary endpoint results of the NELSON trial have not yet been analyzed, the CoD review process will be continued and is until then not known by study arm.

In the Detection And screening of early lung cancer with Novel imaging TEchnology (DANTE) trial, 78% of all the death certificates were cross-checked with hospital records and/or with written reports provided by the attending physicians or by the general practitioner [1]. In event of a doubtful case, the 'death cause review panel' was presented with all available medical records (blinded for patients' assignment to the LDCT or control group) during formal case review sessions to reach a consensus about the CoD. In the Danish Lung Cancer Screening Trial (DLCST), the medical history of the deceased participants was also obtained, and a 'death review board' was formed to establish the cause of death [18]. However, in both trials a protocol about the death review process and information regarding reclassifications after the review off the CoD are not available. So, therefore, no comparison of the death review can be made with the DANTE and DLCST trial.

Strengths of this study are the blinded and uniform review process [12]. This allowed the reviewers to evaluate the medical file without the knowledge whether the participant was randomized in the screen arm or control arm. Furthermore, during the review process, the CEC members and the third reviewer were blinded for what the outcome on the death certificate was. Moreover, in the CEC members and the third reviewer had access to participant's complete medical file.

Potential limitations of the present study relates to the sample size and the selection of subjects of this study. The first 268 deceased participants from whom all medical data was collected were selected. This may have introduced a selection bias of participants with an aggressive lung cancer leading to more assignable death due to lung cancer. This could also explain why

compared to the NLST fewer official death certificates were re-classified. Secondly, a subset of deaths was selected for the CoD review (e.g. participants with a diagnosis of lung cancer or with a notation of lung cancer on the death certificate), with the chance of selection on the likelihood of lung cancer as cause of death. Also, as excluding all the other participants who died of another CoD, it is unknown whether the results of the CoD review process are generalizable for the whole deceased NELSON population. However, reviewing all death cases is very time-consuming and therefore many large cancer screening trials choose to review only deaths in specific cancer patients [6, 19, 20]. Another limitation is that it is not yet allowed to analyze the data by study arm and the pending lung cancer mortality analysis.

In conclusion, for large randomized cancer screening trials it is necessary to review the cancer-specific cause of death because the official death certificate report's varying accuracy (sensitivity and specificity). In the NELSON trial, so far, a high overall sensitivity and specificity of the official death certificates was observed; in other words possible biased related to lung cancer death seem relatively small. Furthermore, it is recommended that the death review of each case should be done by at least two members, independently and in a uniform and blinded matter.

ACKNOWLEDGEMENTS

The authors thank the system controllers R. Faber and F.J.P. Santegoets, and the secretary M.Quak (all from the department of Public Health, Erasmus University Medical Center Rotterdam for their contribution and maintenance of the database. Furthermore, we thank R. Ziengs (University Medical Center Groningen), S. van Amelsvoort-van der Vorst (University Medical Center Utrecht) and M.S.G. den Uijl (Erasmus Medical University Center). Finally, we thank the Dutch Cancer Registry (NKR) and the Statistics Netherlands (CBS) for the data linkages.

SUPPLEMENTARY MATERIALS

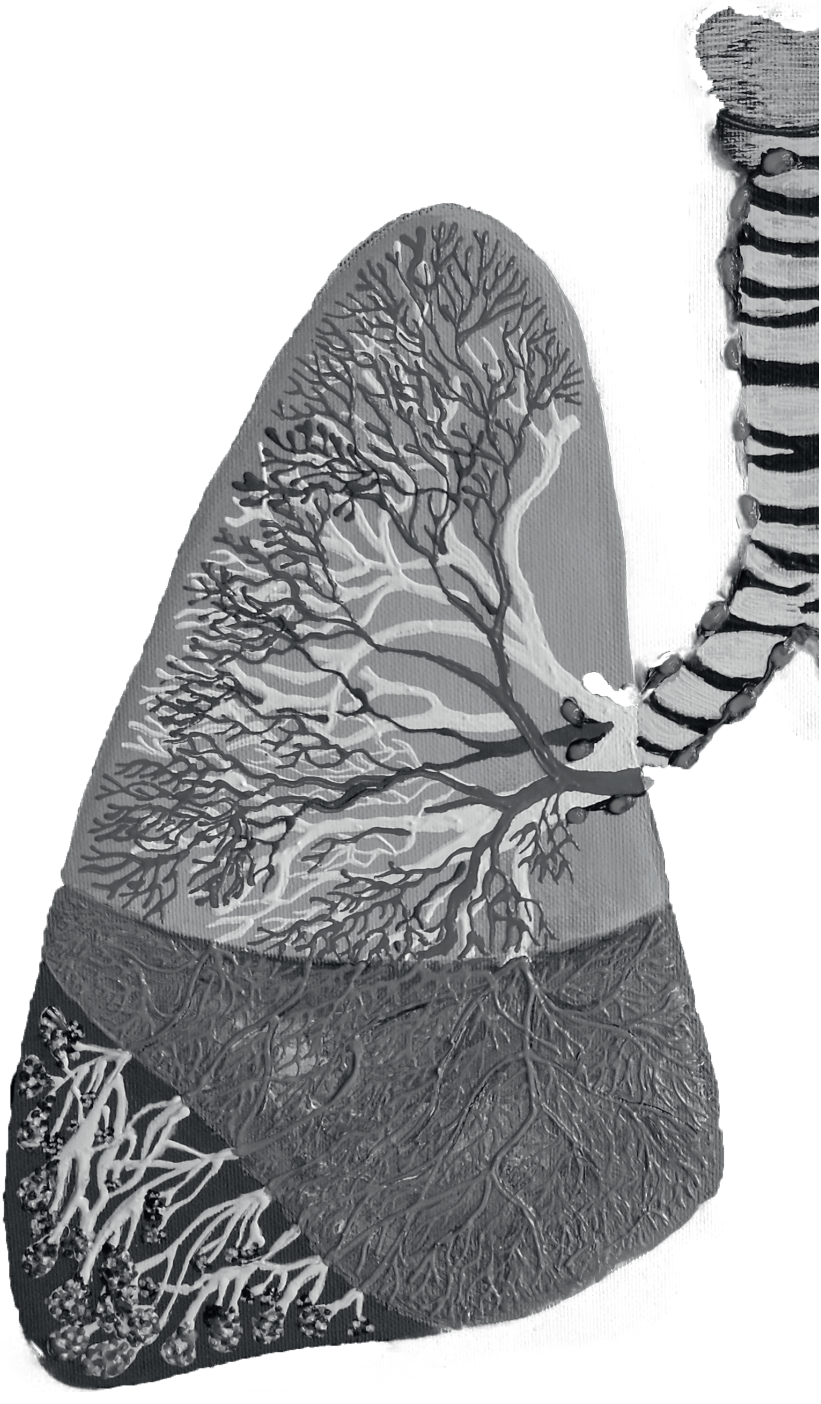
Table 1: classification of the cause of death[12]

Cause of death	Definition
Definitely LC death	Death certainly as a direct result of (second primary) lung cancer, a paraneoplastic syndrome or a diagnostic or therapeutic intervention, including euthanasia and palliative sedation. No clear other cause of death is present
Probable lung cancer Death	Participants with (second primary) lung cancer with evidence of locoregional or distant disease progression or a paraneoplastic syndrome. It is uncertain whether this is the final direct cause of death. No clear other cause of death is present
Possible lung cancer Death	Participants with (second primary) lung cancer with evidence of locoregional or distant disease progression or a paraneoplastic syndrome and one or more coinciding malignancies. It is not possible to determine which malignancy was the primary cause of death
Unlikely lung cancer Death	Participants with (second primary) lung cancer, but without evidence of locoregional or distant disease progression, a paraneoplastic syndrome or death as a result of an intervention for lung cancer. No clear other cause of death is present
Definitely no lung cancer death	The cause of death is definitely not a direct or indirect result from (second primary) lung cancer, a paranoplastic syndrome or an intervention for lung cancer. Another cause of death is present.
Intercurrent death with lung cancer as contributing factor	Only use this option when the cause of death cannot be classified as listed above. The cause of death is definitely not a direct result from (second primary) lung cancer. Another cause of death is present and lung cancer contributed to the death of the patient

REFERENCES

1. Infante M, Cavuto S, Lutman FR, Passera E, Chiarenza M, Chiesa G, et al. Long-Term Follow-up Results of the DANTE Trial, a Randomized Study of Lung Cancer Screening with Spiral Computed Tomography. Am J Respir Crit Care Med. 2015;191(10):1166-75.
2. National Lung Screening Trial Research T, 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.
3. 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 Med. 2009;361(23):2221-9.
4. Williams NJ, Hill EM, Ng SY, Martin RM, Metcalfe C, Donovan JL, et al. Standardisation of information submitted to an endpoint committee for cause of death assignment in a cancer screening trial - lessons learnt from CAP (Cluster randomised triAl of PSA testing for Prostate cancer). BMC Med Res Methodol. 2015;15:6.
5. Otto SJ, van Leeuwen PJ, Hoekstra JW, Merckelbach JW, Blom JH, Schroder FH, et al. Blinded and uniform causes of death verification in cancer screening: a major influence on the outcome of a prostate cancer screening trial? Eur J Cancer. 2010;46(17):3061-7.
6. Marcus PM, Gareen IF, Miller AB, Rosenbaum J, Keating K, Aberle DR, et al. The National Lung Screening Trial's Endpoint Verification Process: determining the cause of death. Contemp Clin Trials. 2011;32(6):834-40.
7. Maudsley G, Williams EM. Death certification by house officers and general practitioners--practice and performance. J Public Health Med. 1993;15(2):192-201.
8. Degani AT, Patel Rm Fau - Smith BE, Smith Be Fau - Grimsley E, Grimsley E. The effect of student training on accuracy of completion of death certificates. (1087-2981 (Electronic)).
9. Aung E, Rao C Fau - Walker S, Walker S. Teaching cause-of-death certification: lessons from international experience. (1469-0756 (Electronic)).
10. Horeweg N, van der Aalst CM, Thunnissen E, Nackaerts K, Weenink C, Groen HJ, 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.
11. Yousaf-Khan U, van der Aalst C, de Jong PA, Heuvelmans M, Scholten E, Lammers JW, et al. Final screening round of the NELSON lung cancer screening trial: the effect of a 2.5-year screening interval. Thorax. 2017;72(1):48-56.
12. Horeweg N, van Klaveren RJ, Groen HJ, Lammers JW, Weenink C, Nackaerts K, et al. Blinded and uniform cause of death verification in a lung cancer CT screening trial. Lung Cancer. 2012;77(3):522-5.
13. Ru Zhao Y, 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.
14. Xu DM, Gietema H, de Koning H, Vernhout R, Nackaerts K, Prokop M, et al. Nodule management protocol of the NELSON randomised lung cancer screening trial. Lung Cancer. 2006;54(2):177-84.
15. De Koning HJ, Blom J Fau - Merkelbach JW, Merkelbach Jw Fau - Raaijmakers R, Raaijmakers R Fau - Verhaegen H, Verhaegen H Fau - Van Vliet P, Van Vliet P Fau - Nelen V, et al. Determining the cause of death in randomized screening trial(s) for prostate cancer. (1464-4096 (Print)).

16. McHugh ML. Interrater reliability: the kappa statistic. (1330-0962 (Print)).
17. Marcus PM, Doria-Rose VP, Gareen IF, Brewer B, Clingan K, Keating K, et al. Did death certificates and a death review process agree on lung cancer cause of death in the National Lung Screening Trial? Clin Trials. 2016;13(4):434-8.
18. 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.
19. Miller AB, Baines CJ, To T, Wall C. Canadian National Breast Screening Study: 2. Breast cancer detection and death rates among women aged 50 to 59 years. Cmaj. 1992;147(10):1477-88.
20. Makinen T, Karhunen P, Aro J, Lahtela J, Maattanen L, Auvinen A. Assessment of causes of death in a prostate cancer screening trial. Int J Cancer. 2008;122(2):413-7.



CHAPTER VII

General discussion

Summary

Samenvatting

Dankwoord

About the author

PhD Portfolio

List of publications

GENERAL DISCUSSION

In the first paragraph, the main results of the following topics are summarized and interpreted: I) the optimization of the NELSON screening rounds; II) the interim stage shift results of the NELSON trial; and III) the cause of death of the NELSON participants. The methodological considerations are discussed in the second paragraph. Finally, the general conclusions and implications for further research are mentioned in the third paragraph.

Paragraph I

main results of the research topics

PART I: THE OPTIMIZATION OF THE NELSON SCREENING ROUNDS

I. Added value of a fourth screening round

Main research question

1. What is the added value of a fourth screening round with an interval of 2.5 years after the previous three screening rounds?

Main results

All participants from the screen group without lung cancer who were still alive in 2009 (to our best knowledge) (n=6,735) received an invitation to participate in the additional fourth Computed Tomography (CT) lung cancer screening round that was initiated from 2009 to 2012 [1]. In total, 80.7% (5,437/6,735) accepted the invitation, of whom 97.1% (5,279/5,437) attended the final screening round.

In total, 5,380 CT scans were performed of which 5,279 were regular round scans and 101 were follow-up scans to assess the volume-doubling time (VDT) of indeterminate nodules. The lung cancer detection rate in the fourth screening round was 0.8% (43/5,380). The positive predictive value (PPV) was 41.0% and the false-positive rate after positive screening was 59.0%.

In this final screening round, 46 screened-detected lung cancers were detected in 43 participants. Of these cancers, 60.9% were diagnosed in stage I, 15.2% in stage II, 10.8% in stage III, and 13.1% in stage IV. Approximately half of the lung cancers in round four were adenocarcinomas, 21.7% were squamous cell carcinomas (SQM), 8.7% bronchoalveolar carcinomas (BAC), 6.5% small cell carcinomas (SCLC), 8.7% non-small cell carcinomas not otherwise specified (NSCLC NOS), and in 4.4% no histological diagnosis was possible.
The adenocarcinomas tended to be more frequently detected in lower stage (stage I-IIIa) compared with SCLC which were mostly diagnosed in stage III/IV (p=0.06). Age (p=0.81), gender (p=0.38), current smoking (p=0.89) and starting age of smoking (p=0.28) had no influence on the stage at time of diagnosis.

In the 2.5 years between the third and fourth screening round, 28 participants were diagnosed with an interval cancer. The majority of these participants (16/28) received the diagnosis in the last six months before the final fourth screening round took place. Participants with an interval cancer in the first 24 months were slightly younger than the participants with an interval cancer in the last six months (64.2 vs. 65.4 years, p=0.05).

Interpretation of the results

Compared to the first three rounds, the participation rate in the final screening round was slightly lower [1-3]. The original NELSON study protocol consisted of only three CT screening rounds [4]. Therefore, it was necessary to obtain a second informed consent form for this additional screening round. Only those participants who were still alive after the third screening round and from whom the updated addresses were known were invited to participate in the additional fourth screening round. Excluded were those diagnosed with lung cancer in the first three rounds, an interval cancer and those who passed away in the meanwhile [2, 3, 5]. Thus, the high non-response rate may have been caused by factors such as lung cancer diagnosis, the traceability of participants (e.g. updated information about the whereabouts could not be retrieved for a small percentage of the study population), and less interest for an additional fourth screening round at the time of dispatching the invitation, which was almost six years after randomisation and about 2.5 years after completing the first three screening rounds. In this study, it was also observed that participants with solely negative screening results were possibly more interested to participate in the additional fourth screening round compared to those with at least one non-negative result (and without lung cancer) in the first three rounds. Moreover, a higher proportion of current smokers at randomisation attended the final screening round compared with former smokers. This may support the idea that lung cancer screening may have an unintended health certificate effect (e.g. a negative screen outcome permits to continue smoking) in some participants [6]. Therefore, a smoking cessation intervention should be integrated with lung cancer screening.

The lung cancer detection, PPV and false-positive (FP) rates did not significantly differ between the four screening rounds [2, 3, 5]. However, the true positive (TP/ FP) ratio tended to improve across the first three rounds from 0.69 to 0.83 and declined in the fourth screening round to 0.69. This was caused by a slightly (non-significant) increased FP rate in the final round compared to the third screening round (from 54.5% to 59.0%). The longer screening interval of 2.5 years may have led to more abnormalities of which the discrepancy between benign and malignant was not obvious. However, a lower proportion of screening in round four led to an indeterminate or positive scan result than screening in round three. This could be due to the NELSON nodule management screening protocol, in which the radiologists were allowed to categorize stable abnormalities across the previous screening rounds as negative in the final screening round [4, 7]. So far, it is unknown if these stable abnormalities have progressed to an interval cancer diagnosis after the final screening round. Another explanation of lower proportion of indeterminate or positive scan results in the final screening might be the high interval cancer rate in the last six months before the fourth screening round took place.

In total 2.0% of the scans across the four screening rounds were positive, which is comparable to the Danish Lung Cancer Screening Trial (DLCST) [8, 9]. The DLCST used an adapted version of the NELSON nodule management protocol. Compared to the National Lung Screening Trial (NLST), the overall positive screening result (2.0% vs. 24.4%) and the false positive rate (59.4% vs. 96.4%) after a positive screening was substantially lower in the NELSON trial [10, 11].

This difference can be explained through the use of a volumetric nodule management protocol in the NELSON trial, in which the volume, the volume growth and VDT of the nodule was determined to evaluate the potential malignancy of the nodule [4, 12], while in the NLST the nodules were classified according to their diameter [13]. Moreover, the NELSON nodule protocol allowed the radiologists to first categorize the nodule into three categories, depending on a set of features of the nodule defined prior to screening as negative, indeterminate or positive. After a follow-up scan, the indeterminate screening result was classified as negative or positive. In contrast, the use of the NLST nodule management protocol resulted in two categories: negative or positive [13]. The NELSON approach in which a third category (indeterminate) and a volumetric nodule approach were used, probably led to less positive scan results and to less false-positive results after a positive screening result [3, 5]. The NELSON cumulative lung cancer detection rate is 3.2%, which is comparable to the DLCST, which used five annual screening rounds [1, 8]. Relative to the NLST, the lung cancer detection rate in the NELSON trial was higher (2.4% vs. 3.2%). However, the NLST only had three annual screening rounds and the NELSON screening protocol consisted of four screening rounds with increasing screening lengths [10, 11].

Previous NELSON analyses [2, 3] showed that the 2-year screening interval after the second round did not lead to significantly more advanced stage lung cancer compared to the 1-year screening interval after the baseline screening [3]. However, the 2.5-year screening interval led to significantly more stage IIIb/IV cancers compared to a 1-year screening interval [1]. Moreover, the interval cancer rate was the largest in the 2.5-year screening interval: 1.47 times higher than in a 2-year screening interval. In the last six months before the fourth screening round took place the interval cancer rate was 1.3 times higher than in the first 24 months after the third screening round, suggesting that a 2.5-year interval is too long.

Relative to the other lung cancer screening trials, the cumulative stage distribution of the screen-detected lung cancers appears to be more favorable in the NELSON trial: more cancers are diagnosed in stage I (69.4%) compared to the NLST (61.6%) and the DLCST (68.1%), and less cancers are staged IIIb/IV (9.8% vs. 20.0% and 15.9%, respectively) [8-11]. These findings should be interpreted with caution, as the 7th edition of the TNM staging system was used in the NELSON trial, whereas the 6th edition was used in the NLST, complicating the comparison. The 7th edition led more cancers to be categorized in a lower stage than according to the 6th edition [14]. On the other hand, fewer females were included in the NELSON trial, while females with lung cancer have a more favorable prognosis and are more often diagnosed with relatively slow-growing cancers [15]. Furthermore, the overdiagnosis rate has yet to be determined in the NELSON trial. In the NLST 18.5% of the screen-detected lung cancers are being reported as overdiagnosis [16].

Conclusion

A 2.5-year screening interval after a third screening round reduces the effect of screening. In the final screening round, significantly more advanced staged lung cancers were detected

as compared to a 1-year screening interval. A similar unfavorable stage distribution of the screen-detected lung cancers was noted when compared to a 2-year screening interval, however not statistically significant. Not only were there more interval cancers diagnosed in the 2.5-year screening interval compared to one-year and two-year screening intervals, but the largest proportion of interval cancers was diagnosed in the last six months of the 2.5-year screening interval.

II. Risk stratification

Main research question

2. Which NELSON subgroups with different risks for detecting lung cancer can be identified based on their previous screening history?

Main results

The probabilities for the screening outcome and risk for detecting lung cancer in the fourth screening round were calculated based on the regular scan results of the first three rounds and of the third scan result only [17]. The regular scan results are the initial screening results of each round, in contrast to follow-up scans, which were performed to determine the growth in order to classify the indeterminate scan result as negative or positive.

Risk stratifications based on the results of the first three screening rounds
Only those participants who were screened in both the first three rounds and as well as in the fourth screening round were included in this analysis. Three subgroups were identified based on the regular scan results of the first three rounds: 1) participants with only negative scan results (n=3,856); 2) participants with at least one indeterminate scan result, but never a positive scan result (n=1,342); and 3) participants with at least one positive scan result (n=81). Participants with only negative scan results were significantly younger than the other two subgroups (57.0 year vs. 58.0 year, $p>0.001$). The third subgroup had smoked slightly more pack-years than the other subgroups (38.7 years vs. 38.0 years, $p=0.02$). The second subgroup (odds ratio (OR) 1.89, $p=0.001$) and the third subgroup (OR 3.77, $p<0.001$) had significantly higher ORs to receive a non-negative (indeterminate or a positive scan result) scan result in the fourth screening round compared to the first subgroup with solely negative scan results so far.

In multivariate analysis, screening history ($p<0.001$) and smoked pack-years ($p=0.02$) remained significant predictors for the regular scan result in the final fourth screening round. No interaction was found between the variables screening history and smoked pack-years ($p=0.89$).

The OR for detecting lung cancer in the final fourth screening round for the second subgroup 2.77 ($p<0.001$) relative to first subgroup with solely negative scan results so far. In the final

screening round, no lung cancer was detected in third subgroup.

In total, 43 participants had a screen-detected lung cancer in round four. So far, 51.2% (22/43) of these participants had solely negative scan results in the previous three screening rounds. In 20 out of these 22 participants, the screen-detected lung cancer arose from a newly detected nodule in the final screening round. The other 21 participants with a screen-detected lung cancer in round four, had in the first three screening rounds at least one indeterminate but never a positive scan results. In 12 out of these 21 participants the lung cancer was detected in a newly detected nodule in the final round.

Participants with only negative scan results in the first three rounds were stratified for their smoked pack-years. Six categories were made: ≤25 years, 26 to 30 years, 31 to 35 years, 36 to 40 years, 41 to 45 years and >45 years. A history of smoked pack-years of up to 45 years led to a lung cancer risk detection in the fourth screening round of between 0.2% and 0.7%. For those who smoked more than 45 pack-years (n=1,091), the lung cancer detection rate in the fourth round was 1.1% (p=0.04).

Risk stratification based on the scan results of the third round only
The probability for detecting lung cancer in the fourth round for the participants with a negative third scan result was 0.6%, which was significantly lower compared to the participants with an indeterminate scan result in round three (3.7%, p<0.001). No lung cancer was detected among those participants with a positive scan result in the third round.

Participants with a screen-detected lung cancer
Across all four screening rounds, 243 out of 7,582 participants were diagnosed with a total of 255 screen-detected lung cancers. Those with a screen-detected lung cancer were older (61.0 years versus 58.0 years, p<0.001) and had smoked more pack-years (44.0 versus 38.0 years, p<0.001) compared with those without a screen-detected lung cancer.

Furthermore, when regular scan results from all four screening rounds were combined, in 28.4% from the screen-detected lung cancers cases had ≥1 indeterminate scan result (but never a positive test result initially) and 71.6% had ≥1 positive scan result (this group also contains those with once a negative or an indeterminate result) before diagnoses of lung cancer.

Interpretation of the results

The probability for receiving a non-negative scan result in the fourth screening round for those with at least one indeterminate (but never a positive scan) result and for those with at least one positive result in the first three screening rounds was higher compared to those with only negative scan results in the first three screening rounds. Furthermore, higher age and more pack-years smoked were both significant predictors for a non-negative scan result in the final screening round. As found in previous analysis of the first three screening rounds [3].

Retrospectively, for those with solely negative scan results in the first three rounds and for those with a negative third round scan result, the risk for detecting lung cancer in round four was lower than 1%. This lung cancer detection risk is approximately the same as across all the screening rounds for the whole screen group: 0.9% in round 1, 0.8% in round 2, 1.1% in round 3 and 0.8% in round 4 [1-3]. In other words, those with previous solely negative scan results and those with a third negative scan result may not require subsequent screening for more than 2.5 years.

On the other hand, almost 90% of the participants in the group with solely negative scan results had a screen-detected lung cancer in the fourth screening round which grew from a nodule first described in the fourth screening round (new nodule). One explanation of evolving lung cancer from a new nodule is the "field cancerization" theory, in which it is assumed that large areas of the bronchial epithelium are affected by smoking, leading to areas with metaplasia and dysplasia which sometimes can turn into cancer. Furthermore, Walter et al. showed that at each screening round (round two and round three) about 5-7% of the screened participants had a new solid nodule in the NELSON trial [18]. This new nodule had a higher probability of malignancy even at a small size compared to baseline nodules: in round two and three a total of 1,222 new solid nodules were registered in 787 participants. In 49 (6%) participants the lung cancer grew from a new solid nodule, which represents 4% of all the new solid nodules. Moreover, new solid nodules with a volume of larger than 27mm^3 had a lung cancer probability of 3.1% to 16.9%. This volume cutoff is smaller than the cut-off of 50mm^3 in the NELSON nodule management protocol [4]. In conclusion, they suggested a more aggressive follow-up strategy in fourth round screen-detected lung nodules than baseline nodules with a short term follow-up for growth assessment. So, the risk for detecting lung cancer after an interval of 2.5 year is 1.1% at the most but on the other hand more lung cancers are detected in a new solid nodule which needs a more aggressive follow-up. Modeling will hopefully provide more insight in the best scenario.

Only the NELSON trial used different screening intervals in one screen group until now. Therefore, comparison with other lung cancer screening trials is difficult. However, in a recent retrospective analysis the NLST did show that participants with a negative baseline scan result had a lower incidence of lung cancer at baseline as well as a lower lung cancer detection rate in the second and third screening round (0.34%), compared to all screened participants (1.0%) [19]. Furthermore, it showed that those with solely negative scan results across the three screening rounds had an even lower lung cancer incidence and mortality rate than those with a negative baseline scan result. This suggests that is it indeed possible to identify subgroups with significantly different lung cancer probabilities based on the screening history.

In conclusion, screening history might be used for further risk stratification of subjects who undergo lung screening and it may lead to a better harm-benefit ratio for a large part of the screened population (e.g. reducing the number of scans needed).

PART II: INTERIM STAGE SHIFT RESULTS IN THE NELSON TRIAL

III. Cancer stage shift and treatment shift

Main research question

3. What is the level of cancer stage and treatment shift between the two study groups?

Main results

The first 100 lung cancers diagnosed in the screen group were compared with the first 100 diagnosed lung cancers in the control group. The lung cancers in the screen group were diagnosed a year earlier on average than the lung cancers diagnosed in the control group (2005 vs. 2006, p<0.001).

More favorable outcomes were observed in the screen group, when comparing the screen and control group: more stage I (59.0% vs. 19.0%) and less stage IV (6.0% vs. 48.0%) lung cancers were found in the screen group. Furthermore, a higher proportion of adenocarcinomas (45.0% vs. 27.0%) and bronchoalveolar carcinomas (BAC; 3.0% vs. 0.0%) were observed in the screen group (p<0.001). Also, significantly lower proportions of squamous cell carcinomas (SQM; 21.0% vs. 25.0%), large cell carcinomas (11.0% vs. 19.0%) and small cell carcinomas (SCLC; 2.0% vs. 18.0%) were observed in the screen group compared to the control group (p<0.001). Stage I cancers stratified for each study group showed that the proportion of stage Ia cancers was the highest for the screen group (91.7% vs. 47.4%, p <0.001) compared to the control group. No differences were found in histology of stage I (p=0.13) lung cancers between the screen and control group.

Screen group lung cancer patients received a (potentially curative) surgical treatment more often regardless of the stage (67.7% vs. 24.5%, p<0.001). Even stratified for stage Ia, screen group participants received more often a (potentially curative) surgical treatment than control group participants (82.7% vs. 50.0%, p=0.06; data not shown). No difference was seen in age at diagnosis, smoking status and co-morbidity between the screen group and control group participants with stage Ia cancer.

Compared to the same Dutch birth cohort with lung cancer, the lung cancers diagnosed in the control group were diagnosed in a slightly earlier stage (stage I or stage II, 29.1% vs. 32.5%, p<0.001). No difference was observed in the first lung cancer treatment between the lung cancer patients of the NELSON control group or the comparable Dutch birth cohort.

Interpretation of the results

CT screening for developing lung cancer in high risk subjects has led to a shift in terms of early detection of lung cancer compared to no screening. Not only were more cancers

detected at an early stage (stage I), but also less advanced staged cancers were found (stage IV) in the screen group. Furthermore, early detection of lung cancer enhanced the probability of a (surgical) treatment that is more likely to be curative in the screen group.

Compared to the NLST, the cancer stage shift in the NELSON trial is substantially more favorable: more stage I cancers (60.0% vs. 50.0%), and less stage IV cancers (6.0% vs. 21.7%) were detected in the NELSON [1-3, 5, 10, 11]. The NELSON trial used the 7th TNM staging system, whereas the NLST used the previous 6th TNM staging system, which might explain part of this difference [14]. On the other hand, in the NELSON trial fewer women were randomized (due to study selection criteria) compared to the NLST, while women are often diagnosed with lung cancer at a lower stage [15]. Furthermore, the NELSON screening protocol consisted of increasing screening intervals between the four screening rounds, while the NLST screening protocol consisted of three annual screening rounds [4, 13]. However, current NELSON analysis includes only the first 100 lung cancers of each study group. Thereby the second screening round was still running in the screen group. Altogether, this sub study showed that it seems that the NELSON screening strategy is at least as capable as the NLST to diagnose lung cancer at a more favorable stage compared to the current setting (no screening) in the Netherlands and Belgium.

Studies have shown that up to 25% of the CT screen-detected lung cancers are relatively slow-growing tumours of which up to 80% are adenocarcinoma or BAC [16, 20, 21]. In the current NELSON analysis, 48.0% of the lung cancers in the screen group are relatively slow-growing tumours, of which 93.8% are adenocarcinomas. Detection of more relatively slow-growing tumours may lead to weakening of the effect of screening and may result in overdiagnosis and overtreatment [16, 22]. On the other hand, screening has also led to more SCLC being diagnosed in a lower stage compared to the control group lung cancers. Further studies are needed to investigate the histology shift in more depth and to determine the proportion of overdiagnosed lung cancers in the NELSON trial.

As lung cancer is mostly diagnosed clinically in an advanced and often metastasized stage, it is surprising to observe that in 19.0% of the lung cancer cases in the control group it was diagnosed in stage I. To be diagnosed in such a low cancer stage, the participant should have been under any kind of supervision of a specialist (e.g. treatment of COPD). However, this aspect was not investigated in this sub study.

Furthermore, screen group participants with a stage Ia lung cancer underwent more often a surgical treatment than control group participants with a stage Ia lung cancer. No difference was seen in age at diagnosis or in co-morbidity (such as cardiovascular diseases or COPD) between these groups of participants.

This sub study showed that clinically diagnosed lung cancer in the comparable general Dutch population was diagnosed slightly at a higher stage than in the NELSON control group participants. One potential explanation is the study eligibility in which subjects with a

history of cancer (lung cancer in the past five years or longer than five years ago but still under treatment, or those with renal cancer, melanoma or breast cancer) or those with a moderate or bad self-reported health who were unable to climb two flight of stairs and those with a CT chest examination less than one year ago before they filled in the NELSON questionnaire were excluded from participation in the NELSON trial [4, 23]. As the actual difference in stage between the NELSON control group and the Dutch comparable birth cohort is small, it is expected that the NELSON trial end results will be generalizable for the target population.

In conclusion, LDCT screening in the NELSON trial has so far led to a substantial favorable shift in cancer stage at time of diagnosis. Implementing of lung cancer screening will lead to higher rate of early stage lung cancers, and subsequently a shift in treatment options from mainly advanced disease stage treatment (e.g. palliative) to curative treatment (mainly surgery).

PART III: THE CAUSE OF DEATH OF THE NELSON PARTICIPANTS

IV. Healthy participant effect and mortality profile

Main research question

4. What are the differences in characteristics and mortality profile of NELSON participants and eligible non-responders?

Main results

A total of 7,435 Dutch control group participants and 13,661 Dutch eligible non-responders were compared regarding background variables such as age, educational level, and smoking history [24]. Non-responders were those who met the eligibility criteria of the NELSON trial but who did not provide informed consent and who did not participate in the trial.

The NELSON participants were younger ($p<0.001$), more often male ($p<0.001$), had a better self-reported general health ($p=0.002$), had a higher level of physical exercise ($p<0.001$), had a higher level of education ($p<0.001$), and were more often former smokers ($p<0.001$) as compared to the Dutch eligible non-responders.

Smoking duration was lower among participants ($p<0.001$), whereas numbers of cigarettes smoked per day was higher among participants ($p<0.001$). Participants started smoking at a younger age ($p<0.001$) and were more willing to quit smoking than eligible non-responders ($p<0.001$). Among current smokers, participants were more often in an advanced stage- according to the stages of change- to quit smoking compared with eligible non-responders ($p<0.001$). Participants reported significantly more smoking-related symptoms ($p = 0.04$) and had undergone a pulmonary function test more often ($p<0.001$). However, no differences were seen in the self-reported outcome of these pulmonary function tests ($p=0.28$).

Until January 2013, significantly more Dutch eligible non-responders deceased than NELSON control group participants (11.2% vs. 9.1%, $p<0.001$). The mortality rate due to all types of cancer (7.59 vs. 6.32, $p=0.002$), cardiovascular diseases (3.32 vs. 2.05, $p<0.001$), respiratory diseases (0.73 vs. 4.4, $p=0.018$), and non-cancerous diseases other than cardiovascular and respiratory (2.19 vs. 1.30, $p<0.001$) was highest in the Dutch eligible non-responders. The proportion of deaths due to cancer was higher among participants (62.4% vs. 54.9%).

Interpretation of the results

This study determined whether the NELSON trial end results (mortality results) are generalizable to the target Dutch population for lung cancer screening.

In line with previous cancer screening trials [25-34], a small healthy participant effect in the NELSON trial was observed: the NELSON participants were significantly slightly younger, had a better self-reported health, were more psychically active, and were higher educated compared with eligible non-responders.

In contrast to the DLCST and the ITALUNG, less current smokers participated in the NELSON trial [26, 28]. On this matter, the DLCST reported that despite more active smoking, DLCST participants were more willing to quit than non-participants (a representative sample from the Danish population), suggesting that smokers who are motivated to quit smoking are more inclined to volunteer in a screening trial [28]. This is supported by our finding that among current smokers, participants were more often in an advanced stage -according to the staged of change- to quit smoking compared with eligible non-responders. On the other hand, other studies have shown that active smoking is a barrier to undergo screening for lung cancer [30, 35]. Moreover, those with a lower socio-economic status and a lower educational level are more likely to smoke. Future studies are needed to further assess the interactions between socioeconomic status and active smoking on attitudes toward lung cancer screening

Eligible non-responders had a higher all-cause mortality and mortality due to four other mortality classifications. However, the relative proportion of subjects that died due to all types of cancer was higher among participants. This might be explained by a higher alcohol abuse among participants, which is associated with a higher relative death from cancer [36]. Moreover, higher alcohol consumption is also related with a higher socioeconomic status and a higher educational level of achievement [37]. Another explanation could be that the participants reported more smoking-related symptoms, (e.g. dyspnea and hemoptysis) which may have led to more doctor consults. Hence, may leading to more participants to quit smoking and could have facilitated the detection and early treatment of cancer, cardiovascular, and respiratory diseases. However, the younger age, healthier lifestyle, a higher educational level and a better self-reported health among participants may have led to a higher contribution to the difference in mortality profiles between the NELSON participants and eligible non-responders.

In conclusion, the NELSON study participants are somewhat younger, healthier, higher educated and more willing to participate in a screening program than the eligible non-responders as expected . These differences have influenced the mortality outcomes of both groups. However, the differences are modest and therefore it seems unlikely that these differences will influence the overall generalizability of the main results of the NELSON trial towards the target population.

PART IV: CAUSE OF DEATH VERIFICATION PROCESS

Main research question

5. What is the outcome of the Cause of Death verification process in the NELSON trial, and how does it relate to the official death certificates?

Main results

The clinical expert committee (CEC) of the NELSON trial consists of mainly two members who independently of each other reviewed 266 medical files of Dutch deceased NELSON trial participants [38]. A decisional-flow chart and a detailed list of criteria were used to make a decision in a uniform matter [39].

Based on six possible outcomes (defining the graduation of certainty that lung cancer was the primary cause of death (CoD)) an agreement between the two members was reached in 71.1%. Divided in a lung cancer specific death (definitely or probably lung cancer death) and in other CoD (possible, unlikely or definitely no lung cancer death, and intercurrent death with lung cancer as a contributing factor) agreement between the two members was reached in 86.1%. The latter corresponds with a Cohen's kappa of 0.57 (95%CI, 0.45-0.69, $p<0.001$). Cases without agreement were discussed in a consensual meeting. No third reviewer was necessary to reach consensus.

Compared to a previously performed pilot study, in which two members of the NELSON trial tested the utility of the decisional flow-chart, seven out of 49 cases showed no agreement in the CoD [39]. These seven cases needed a re-evaluation by the CEC and the clinical epidemiologist from the pilot study.

Comparison of the CoD by the CEC and the CoD according to the official death certificates showed an overall sensitivity and specificity of the death certificates of 92.6% and 98.8%, respectively. CoD review by a CEC resulted in a reclassification of 12.2% of the cases. In 15 out of these 32 (46.9%) cases, the CEC reclassified the CoD to a lung cancer specific death instead of 'other CoD' as noted on the official death certificate. In the other 17 cases (53.5%), the CEC reviewed the CoD as 'another CoD', while according to the official death certificate the UCD was lung cancer death.

Interpretation of the results

In the NELSON trial all medical files of the selected deceased participants were independently reviewed by two members (in a blinded and uniform manner). A weak to moderate agreement between the two members was observed, which slightly increased with the number of cases they evaluated. This underlies the need that each death should be reviewed by at least two independent experts. The results are in line with the NLST, in which 42% of the deaths were

reviewed and a sensitivity of 91% and a specificity of 97% of the official death certificates were reported [40, 41]. However, in their death review process initially only one member reviewed the medical files. Only those cases without concordance with the official death certificate were reviewed by a second member.

Furthermore, less deaths were re-classified in the NELSON trial compared to the NLST (12.2% vs. 26.0%) [40, 41]. The NLST showed a higher reclassification of deaths to lung cancer specific deaths than the NELSON trial (22.0% vs. 6.5%). This may suggest that in the Netherlands a higher accuracy of clinicians is seen regarding reporting deaths related to lung cancer or due to other causes of death. However, in the NELSON trial, so far, fewer cases have been reviewed. Moreover, the NLST showed that the mortality benefit of screening did not significantly change with reviewing the CoD [41]. As the primary endpoint results of the NELSON trial have yet to be analyzed, the CoD review process will be continued and is until then not known by study arm.

Comparison with other lung cancer screening trials (DLCST and DANTE) was not possible, as in both trials a protocol about the death review process and information regarding reclassification after the review of the CoD was not available [8, 42].

In conclusion, for large randomized cancer screening trials it is necessary to review the cancer-specific cause of death due to the official death certificate reports varying accuracy (sensitivity and specificity). In the NELSON trial, so far, the results indicated that possible biases related to lung cancer death seem relatively small. Furthermore, it is recommended that the death review of each case should be done by at least two members, independently and in a uniform and blinded manner.

Paragraph III

general conclusions and recommendations

Strengths

Population-based randomized-controlled trial

One of the main strengths of this study is the population-based randomized-controlled setting of the trial, in which large numbers of participants are randomly allocated to either the screen or control group [4]. The use of this study design is considered the most preferable one [43]. Only in a randomized-controlled setting, disease-specific mortality between the screened and the unscreened population can be compared, without lead-time, length-time and overdiagnosis biasing the comparison between the two groups. Moreover, randomisation is assumed to lead to a comparable distribution of both known and unknown factors in the two groups which are being compared. Observational studies are prone to different biases and thus to erroneous study results [44].

Secondly, in volunteer-based recruitment, volunteers are unlikely to be representative of the target population. Moreover, those not motivated to volunteer may be particularly important to reach. A population-based recruitment method has therefore been used in the NELSON trial [4, 6, 23]. It allows the results to be generalizable to the whole target population.

Comparison with no intervention

Another major strength of this study is that the control group never received any form of screening or contact with the screening site [4]. For example in the Danish Lung Cancer Screening Trial (DLCST) the control group received annually spirometry and attended yearly the screening site at which the smoking status was determined along with other assessments [9]. The NELSON control group received usual care, which consists of no regular chest examinations by a chest radiography or CT scan. This allows an unbiased lung cancer incidence comparison between the two study groups. However, as mentioned in this thesis, the NELSON control group participants were for example younger, healthier and were more often former smokers than the eligible non-responders. This might overestimate the impact of screening.

The nodule management and screening protocol

The volumetric-based nodule management with the inclusion of an indeterminate test result in the NELSON trial has led to less false-positive results compared to the other lung cancer screening trials, which used a diameter-based nodule management [1, 5, 8, 10-12, 45]. Less false-positive results means less invasive diagnostic procedures necessary to distinguish between a benign and malignant nodule. Moreover, it has led to identification of subgroups with different risks for detecting lung cancer in a subsequent screening round [3, 17]. These findings are useful to optimize the harm-benefit ratio of a lung cancer screening programme. For example, not all eligibles for a CT lung cancer screening need to undergo it on yearly base. Some may have a better harm-benefit ratio if they undergo screening with a longer screening interval.

The NELSON trial is the only lung cancer screening trial with the use of increasing screening intervals in one screen group [4, 12]. Therefore, it has a unique opportunity to investigate the test characteristics and cost-effectiveness of each screening round over time [1, 2, 5]. Consequently, this information is also useful to optimize the harm-benefit ratio of a lung cancer screening programme.

The follow-up period

The NELSON trial intends to have a long follow-up period of ten years before the mortality analysis will follow [4]. To observe a lung cancer incidence difference in both groups over time, and to assess the excess incidence rate and the overdiagnosis ratio, it is necessary to have a long follow-up period [16, 46].

Finally, the cause of death (CoD) review process is being performed in a blinded and uniform matter [39]. This allows the reviewers to evaluate the medical files without the knowledge whether the participant was randomized in the screen group or control group. During the review process all relevant medical data was accessible online by all members of the review committee. Moreover, during the review process the review committee had no access to the official death certificates, allowing them not to be biased.

Limitations

Study protocol

Regardless of the important strengths of this study, there are also some limitations. Lung cancer screening has been argued as a teachable moment for smoking cessation. At baseline some smoking related questions were assessed from the whole study population. The impact of screening on smoking abstinence was investigated previously by follow-up questionnaires after two and four years of the start of the study [6, 47]. The results suggested that lung cancer screening was a potential teachable moment for smoking cessation. However, this was only assessed in a small sample and there is a remaining concern that participants may experience a false feeling of reassurance after lung cancer screening.

In the first recruitment only males were invited to participate in the NELSON trial resulting in less females being included in the NELSON trial compared to some other lung cancer screening trials [4, 9, 13]. Our results are therefore mostly applicable to males in general, although post-hoc analysis of the NLST showed only weak evidence for differences in the impact of lung cancer screening by sex [48].

Self-reported data was used to recruit study participants. Self-reported data is a commonly used method for data collection, however the use of self-reported data is often under discussion at is depends on how accurately the (retrospective) data is reported by the subject

[49]. Furthermore, questions about the socioeconomic class were limited, and no questions were asked about the ethnic background or psychosocial profile as it is known that among the non-Caucasian subjects the participation rate in screening programs is lower [50-53].

The participation rate

As expected in a long running screening trial, the participation rate in the screen group dropped from the first to the final fourth screening round. One explanation therefore is that a large number of participants was not traceable by the end of the third screening round. During the study course, the last update of the actual addresses of the study participants was held in 2006 during an intensive period with retrieving updated information from approximately 100 communities, while the end of the third round was in 2009. Unfortunately, at the moment of dispatching the second informed consent and an invitation letter to participate in the added fourth screening round, it was not possible to have direct access to their actual addresses by a linkage with the GBA administration. Therefore, it is unknown whether these non-responders were willing to participate or not.

Eligible non-responders

Next, only aggregated data was available from Statistics Netherlands for the eligible non-responders group to compare it with the mortality outcome of the NELSON control group participants. Therefore, it was not possible to perform multi-variate analyses in detail. However, mortality profile differences between the NELSON control group participants and eligible non-responders are modest and will therefore probably not influence the generalizability of the NELSON trial end results.

Cause of death verification process

Currently, only the medical files of deceased NELSON study participants with lung cancer are being reviewed concerning their cause of death [39]. Medical files of deceased participants without lung cancer are not being reviewed by the clinical expert committee. Therefore, it is unknown whether the cause of death process outcome is generalizable to the whole study population. On the other hand, cancer-specific death verification is a commonly used approach as verifying all death cases is very time-consuming [40, 54].

Pending mortality analysis

The NELSON screening strategy is at least as capable as the NLST to diagnose lung cancer at a more favorable stage compared to the control group with no intervention. Moreover, it has led to a shift in treatment options from mainly advanced disease stage treatment to curative treatment. However, at this moment it is impossible to investigate the correlation between stage shift, treatment shift and lung cancer mortality reduction.

General conclusions from this thesis

- A 2.5-year screening interval after a third screening round seems to reduce the possible effect of lung cancer screening. The screen-detected lung cancers were significantly more often diagnosed in an advanced disease stage compared to a one-year screening interval.

- The screening history may be used as a risk stratification tool to refine the screening protocol for subgroups of people who undergo screening.

- Low-dose Computed Tomography screening in the NELSON trial has led to a favorable cancer stage shift between the screened group and unscreened group. Subsequently, screening has also led to more treatment (e.g. curative) options for lung cancer.

- A small healthy participant effect has been observed in the NELSON trial. Although the actual numerical differences were minimal. Therefore, the study results seem applicable to the target population.

- Comparison of the independently reviewed blinded medical files (for patients' identity and study group) of the deceased NELSON participants with the official death certificates showed that possible biases related to lung cancer death seem relatively small. It is recommended that each death should be reviewed independently by at least two clinical experts.

General recommendations for research and implementation

- Future research is recommended to investigate how to reach the population who is (not) willing to participate, as it necessary to increase the participation of those who are expected to benefit the most from lung cancer screening. Moreover, the target population is a unique population consisting of long-term smokers who have increased risk to have smoking-related comorbidities and may experience stigma and battle nicotine addiction.

- Recruitment of a high risk group based on smoking history requires the willingness of participants to share information of their smoking history. Future research should focus on obtaining this crucial information (e.g. web based registration system or personal contact).

- Screening history seems to be a useful risk stratification tool; however this should be investigated further, for example in the currently implemented lung cancer screening in the United States. Moreover, it is unknown if a longer screening-interval for certain subgroups is cost-effective.

- The National Lung Screening Trial (NLST) and the NELSON trial used the two most significant predictors of lung cancer risk: age and smoking history. However, addition of other epidemiological risk factors, such as used in the UK Lung Cancer Screening pilot trial (UKLS) (a 5-year lung cancer risk of ≥5% based on the LLP_{v2}) should be tested in a pilot study, as this may increase the cost-effectiveness of a lung cancer screening programme.

- The lung cancer detection rates and the stage distribution of the screen-detected lung cancers in the NELSON trial are at least as favorable as in the NLST. However, the mortality analyses of the NELSON trial are pending. These results are necessary to verify the mortality reduction as shown by the NLST.

- Pooling data of the European lung cancer screening trial is suggested to obtain more information about the cost-effectiveness of several subgroups.

- Screening leads to more cancers being detected at an early stage. This opens the field and need for more development of minimal invasive treatment options in early lung cancer.

- Smoking cessation is not only the most effective primary prevention method for lung cancer, but it is also for cardiovascular diseases and Chronic Obstructive Pulmonary Disease (COPD). When implementing lung cancer screening, smoking cessation counseling or a smoking cessation program should be an integrated part of it.

- An evaluation should be done to estimate the demand of specialists related to a lung cancer screening programme (e.g. general practitioners, radiologists, pulmonologists and pathologists) in Europe. Furthermore, structured data collection is necessary for a successful screening program.

REFERENCES

1. Yousaf-Khan U, van der Aalst C, de Jong PA, Heuvelmans M, Scholten E, Lammers JW, et al. Final screening round of the NELSON lung cancer screening trial: the effect of a 2.5-year screening interval. Thorax. 2017;72(1):48-56.
2. Horeweg N, van der Aalst CM, Thunnissen E, Nackaerts K, Weenink C, Groen HJ, 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.
3. Horeweg N, van der Aalst CM, Vliegenthart R, Zhao Y, Xie X, Scholten ET, et al. Volumetric computed tomography screening for lung cancer: three rounds of the NELSON trial. Eur Respir J. 2013;42(6):1659-67.
4. 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 Med. 2009;361(23):2221-9.
5. Horeweg N, Scholten ET, de Jong PA, van der Aalst CM, Weenink C, Lammers JW, et al. Detection of lung cancer through low-dose CT screening (NELSON): a prespecified analysis of screening test performance and interval cancers. Lancet Oncol. 2014;15(12):1342-50.
6. van der Aalst CM, van Klaveren RJ, de Koning HJ. Does participation to screening unintentionally influence lifestyle behaviour and thus lifestyle-related morbidity? Best Pract Res Clin Gastroenterol. 2010;24(4):465-78.
7. Wang Y, van Klaveren RJ, van der Zaag-Loonen HJ, de Bock GH, Gietema HA, Xu DM, et al. Effect of nodule characteristics on variability of semiautomated volume measurements in pulmonary nodules detected in a lung cancer screening program. Radiology. 2008;248(2):625-31.
8. 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.
9. 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. J Thorac Oncol. 2009;4(5):608-14.
10. National Lung Screening Trial Research T, 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.
11. Aberle DR, DeMello S Fau - Berg CD, Berg Cd Fau - Black WC, Black Wc Fau - Brewer B, Brewer B Fau - Church TR, Church Tr Fau - Clingan KL, et al. Results of the two incidence screenings in the National Lung Screening Trial. (1533-4406 (Electronic)).
12. Xu DM, Gietema H, de Koning H, Vernhout R, Nackaerts K, Prokop M, et al. Nodule management protocol of the NELSON randomised lung cancer screening trial. Lung Cancer. 2006;54(2):177-84.
13. Aberle Dr Fau - Berg CD, Berg Cd Fau - Black WC, Black Wc Fau - Church TR, Church Tr Fau - Fagerstrom RM, Fagerstrom Rm Fau - Galen B, Galen B Fau - Gareen IF, et al. The National Lung Screening Trial: overview and study design. (1527-1315 (Electronic)).
14. Marshall HM, Leong SC, Bowman RV, Yang IA, Fong KM. The science behind the 7th edition Tumour, Node, Metastasis staging system for lung cancer. Respirology. 2012;17(2):247-60.

15. Kligerman S, White C. Epidemiology of lung cancer in women: risk factors, survival, and screening. AJR Am J Roentgenol. 2011;196(2):287-95.
16. 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.
17. Yousaf-Khan U, van der Aalst C, de Jong PA, Heuvelmans M, Scholten E, Walter J, et al. Risk stratification based on screening history: the NELSON lung cancer screening study. Thorax. 2017.
18. Walter JE, Heuvelmans MA, de Jong PA, Vliegenthart R, van Ooijen PM, Peters RB, et al. Occurrence and lung cancer probability of new solid nodules at incidence screening with low-dose CT: analysis of data from the randomised, controlled NELSON trial. LID - S1470-2045(16)30069-9 [pii] LID - 10.1016/S1470-2045(16)30069-9 [doi]. (1474-5488 (Electronic)).
19. Patz EF, Jr., Greco E, Gatsonis C, Pinsky P, Kramer BS, Aberle DR. Lung cancer incidence and mortality in National Lung Screening Trial participants who underwent low-dose CT prevalence screening: a retrospective cohort analysis of a randomised, multicentre, diagnostic screening trial. Lancet Oncol.May;17(5)(1474-5488 (Electronic)):590-9.
20. 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.
21. Wang Z, Hu Y, Wang Y, Han W, Wang L, Xue F, et al. Can CT Screening Give Rise to a Beneficial Stage Shift in Lung Cancer Patients? Systematic Review and Meta-Analysis. PLoS One. 2016;11(10):e0164416.
22. Han SS, Ten Haaf K, Hazelton WD, Munshi VN, Jeon J, Erdogan SA, et al. The impact of overdiagnosis on the selection of efficient lung cancer screening strategies. Int J Cancer. 2017;140(11):2436-43.
23. van Iersel CA, de Koning HJ, Draisma G, Mali WP, 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.
24. Yousaf-Khan U, Horeweg N, van der Aalst C, Ten Haaf K, Oudkerk M, de Koning H. Baseline Characteristics and Mortality Outcomes of Control Group Participants and Eligible Non-Responders in the NELSON Lung Cancer Screening Study. J Thorac Oncol. 2015;10(5):747-53.
25. Pinsky PF, Miller A, Kramer BS, Church T, Reding D, Prorok P, et al. Evidence of a healthy volunteer effect in the prostate, lung, colorectal, and ovarian cancer screening trial. Am J Epidemiol. 2007;165(8):874-81.
26. Dominioni L, Rotolo N, Poli A, Paolucci M, Sessa F, D'Ambrosio V, et al. Self-selection effects in smokers attending lung cancer screening: a 9.5-year population-based cohort study in Varese, Italy. J Thorac Oncol. 2010;5(4):428-35.
27. van der Aalst CM, van Iersel CA, van Klaveren RJ, Frenken FJ, 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.
28. Hestbech MS, Siersma V, Dirksen A, Pedersen JH, Brodersen J. Participation bias in a randomised trial of screening for lung cancer. Lung Cancer. 2011;73(3):325-31.

29. National Lung Screening Trial Research T, Aberle DR, Adams AM, Berg CD, Clapp JD, Clingan KL, et al. Baseline characteristics of participants in the randomized national lung screening trial. J Natl Cancer Inst. 2010;102(23):1771-9.
30. Silvestri GA, Nietert PJ, Zoller J, Carter C, Bradford D. Attitudes towards screening for lung cancer among smokers and their non-smoking counterparts. Thorax. 2007;62(2):126-30.
31. McCaffery K, Wardle J, Nadel M, Atkin W. Socioeconomic variation in participation in colorectal cancer screening. J Med Screen. 2002;9(3):104-8.
32. Miller AB, Baines CJ, To T, Wall C. Canadian National Breast Screening Study: 2. Breast cancer detection and death rates among women aged 50 to 59 years. Cmaj. 1992;147(10):1477-88.
33. Shapiro JA, Seeff LC, Nadel MR. Colorectal cancer-screening tests and associated health behaviors. Am J Prev Med. 2001;21(2):132-7.
34. Sutton S, Wardle J, Taylor T, McCaffery K, Williamson S, Edwards R, et al. Predictors of attendance in the United Kingdom flexible sigmoidoscopy screening trial. J Med Screen. 2000;7(2):99-104.
35. 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.
36. Gronbaek M, Becker U, Johansen D, Gottschau A, Schnohr P, Hein HO, et al. Type of alcohol consumed and mortality from all causes, coronary heart disease, and cancer. Ann Intern Med. 2000;133(6):411-9.
37. Keyes KM, Hasin DS. Socio-economic status and problem alcohol use: the positive relationship between income and the DSM-IV alcohol abuse diagnosis. Addiction. 2008;103(7):1120-30.
38. Yousaf-Khan AU, van der Aalst CM, Aerts JGJV, den Bakker MA, de Koning HJ.
39. Horeweg N, van Klaveren RJ, Groen HJ, Lammers JW, Weenink C, Nackaerts K, et al. Blinded and uniform cause of death verification in a lung cancer CT screening trial. Lung Cancer. 2012;77(3):522-5.
40. Marcus PM, Gareen IF, Miller AB, Rosenbaum J, Keating K, Aberle DR, et al. The National Lung Screening Trial's Endpoint Verification Process: determining the cause of death. Contemp Clin Trials. 2011;32(6):834-40.
41. Marcus PM, Doria-Rose VP, Gareen IF, Brewer B, Clingan K, Keating K, et al. Did death certificates and a death review process agree on lung cancer cause of death in the National Lung Screening Trial? Clin Trials. 2016;13(4):434-8.
42. Infante M, Cavuto S, Lutman FR, Passera E, Chiarenza M, Chiesa G, et al. Long-Term Follow-up Results of the DANTE Trial, a Randomized Study of Lung Cancer Screening with Spiral Computed Tomography. Am J Respir Crit Care Med. 2015;191(10):1166-75.
43. Harbour R, Miller J. A new system for grading recommendations in evidence based guidelines. Bmj. 2001;323(7308):334-6.
44. Jepsen P, Johnsen SP, Gillman MW, Sorensen HT. Interpretation of observational studies. Heart. 2004;90(8):956-60.
45. Horeweg N, van Rosmalen J, Heuvelmans MA, van der Aalst CM, Vliegenthart R, Scholten ET, et al. Lung cancer probability in patients with CT-detected pulmonary nodules: a prespecified analysis of data from the NELSON trial of low-dose CT screening. Lancet Oncol. 2014;15(12):1332-41.

46. Ten Haaf K, de Koning HJ. Overdiagnosis in lung cancer screening: why modelling is essential. J Epidemiol Community Health. 2015;69(11):1035-9.
47. 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.
48. 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.
49. Connor Gorber S, Schofield-Hurwitz S, Hardt J, Levasseur G, Tremblay M. The accuracy of self-reported smoking: a systematic review of the relationship between self-reported and cotinine-assessed smoking status. Nicotine Tob Res. 2009;11(1):12-24.
50. Rutten LF, Hesse BW, Moser RP, McCaul KD, Rothman AJ. Public perceptions of cancer prevention, screening, and survival: comparison with state-of-science evidence for colon, skin, and lung cancer. J Cancer Educ. 2009;24(1):40-8.
51. Sin MK. Lung cancer disparities and African-Americans. Public Health Nurs. 2017.
52. Ryan BM. Differential eligibility of African American and European American lung cancer cases using LDCT screening guidelines. BMJ Open Respir Res. 2016;3(1):e000166.
53. Fiscella K, Winters P, Farah S, Sanders M, Mohile SG. Do Lung Cancer Eligibility Criteria Align with Risk among Blacks and Hispanics? PLoS One. 2015;10(11):e0143789.
54. Makinen T, Karhunen P, Aro J, Lahtela J, Maattanen L, Auvinen A. Assessment of causes of death in a prostate cancer screening trial. Int J Cancer. 2008;122(2):413-7.

SUMMARY

Lung cancer is the leading cause of cancer death among males and second leading cause of cancer death among females worldwide. The most important risk factor for lung cancer is smoking, causing approximately 85% of all lung cancer cases. Smoking cessation or refraining from smoking is the most effective way to prevent lung cancer. Although the prevalence of smoking has decreased, approximately a quarter of the Dutch population above 18 years of age is a current smoker. Moreover, half of the lung cancer cases are currently diagnosed in former smokers. Clinically, lung cancer is often diagnosed in an advanced stage with a five-year lung cancer survival of approximately 15%. This underlines the need for secondary prevention for lung cancer: lung cancer screening. The United States Preventive Services Task Force (USPSTF) requested an independent review and an investigation of the long-term benefit and harms of different lung cancer screening policies. On the basis of this review and modelled screening policies, lung cancer screening has been implemented in the United States. However, in Europe there is a consensus to wait for the mortality results of the NELSON trial, the largest European lung cancer screening trial.

Part I: the optimization of the NELSON screening rounds

The aim of the first research question was to investigate the added value of a fourth screening round with an interval of 2.5 years after the previous three screening rounds (**Chapter II**). Lung cancers diagnosed in the final screening round were compared with the lung cancers diagnosed in the second screening round (with a one-year screening interval) and with the lung cancers diagnosed in the third screening round (with a two-year screening interval). After a 2.5-year screening interval, significantly more lung cancers were diagnosed in an advanced stage compared with a 1-year screening interval. A similar unfavorable stage distribution of lung cancers was observed compared to a 2-year screening interval, though statistically not significant. Furthermore, a higher proportion of interval cancer was diagnosed in the 2.5-year screening interval than in the 1-year and 2-year screening interval. Moreover, the largest proportion of interval cancers in the 2.5-year screening interval was diagnosed in the last six months before the final screening round took place.

In the next chapter (**Chapter III**), it was investigated whether subgroups with different risk for a non-negative screening result and/or lung cancer detection in the fourth screening round could be identified based on their test results in the previous screening rounds. The probability of a non-negative result in the fourth screening round was the lowest for those with solely negative results, compared to those with at least one indeterminate (but never a positive) or at least one positive result in the first three rounds. Furthermore, those subjects with previous solely negative results and those with a third negative scan result may not require screening for more than 2.5 years, as the lung cancer detection rates in these groups were <1% in the fourth screening round. The use of previous screening test results to refine the screening protocol may lead to a better harm-and-benefit ratio for a large part of the screened population (such as reducing the number of scans needed).

Part II: interim stage shift results in the NELSON trial

In **Chapter IV**, the level of cancer stage and treatment shift between the two study groups were investigated. The first 100 lung cancers diagnosed in both study groups were compared. Low-dose CT screening has led to a shift in terms of early detection of lung cancer compared with no screening. Not only were more stage I cancers diagnosed in the screen group, but also fewer stage IV cancers were diagnosed as compared with the control group. Subsequently, this stage shift has enhanced the probability of a curative treatment. The NELSON screening strategy is at least as capable as the NLST screening strategy to diagnose lung cancer at a more favorable stage compared to no screening. Yet it is unknown whether significant lung cancer and all-cause mortality reductions can be achieved amongst the NELSON study participants.

Part III: the cause of death of the NELSON participants

In the first section of this part (**Chapter V**), the differences in baseline characteristics and mortality outcome of the NELSON control group participants with the eligible non-responders for the NELSON trial were investigated. As expected, the NELSON study participants were somewhat younger, healthier, higher educated and more willing to participate in a screening program than the eligible non-responders. These differences influenced the cause of death of both groups slightly: the eligible non-responders had a higher all-cause mortality and mortality due to four other cause of death (cancer-related death, cardiovascular-related death, respiratory-disease related death, and other-cause related death). However, actual numerical differences were small and investigated subjects were large, and therefore it seems unlikely that it will influence the generalisability of the trial results.

In the last section of this part (**Chapter VI**) the process of the verification of the Cause of Death (CoD) of the deceased NELSON study participants as well as the comparability with the official death certificates as provided by Statistics Netherlands were investigated. After an independent and uniform review of the blinded medical files of the deceased lung cancer patients, a weak to moderate agreement was observed between the two members. The agreement increased with the number of cases evaluated, which underlines the need to review all relevant death cases independently by at least two experts. Comparison of the CoD review process outcome with the official death certificates showed an overall sensitivity and specificity of the death certificates of 92.6% and 98.8%, respectively. In 12.2% of the cases the CoD was reclassified by the review committee, in which 46.9% of the cases the CoD was reclassified to a lung cancer specific death. In the NELSON trial, the results indicated that possible biases in CoD related to lung cancer seem relatively small.

DISCUSSION

In **Chapter VII**, the main results were summarized and interpreted. Furthermore, the methodological considerations are discussed in the second part of **Chapter VII** and in the last paragraph, the general conclusion draw from this thesis and the implication for further research are presented.

In the first part of this thesis, the NELSON trial results showed that a 2.5-year screening interval might reduce the effect of screening compared to a 1-year and a 2-year screening interval; more advanced stage lung cancers and more interval cancers were diagnosed during the 2.5-year screening interval. Furthermore, three NELSON subgroups were identified with significantly different risks for a non-negative (an indeterminate or a positive) scan result and a different risk for detecting lung cancer in the final screening round. These risks were based upon their previous screening history during the first three screening rounds. In line with these results, the NLST showed in a retrospective analysis that it is indeed possible to identify subgroups with significantly different lung cancer probabilities based on their screening history. Therefore, screening test results may be a useful risk stratification tool to refine the screening protocol.

In the second part of this thesis, a substantial shift in cancer stage at time of diagnosis was shown as a result of low-dose CT screening in the NELSON trial so far. Implementation of CT lung cancer screening will lead to higher rates of early stage lung cancers, and subsequently a shift in treatment options from palliative to curative (mainly surgery). Thus far, it is unknown whether these findings indicate sufficient lung cancer and all-cause mortality reduction amongst NELSON study participants.

As expected, a small healthy participant effect was observed in the NELSON trial. This may have influenced the mortality outcome of the study participants compared with the eligible non-responders (those who were eligible to participate in the trial but did not participate or did not fill in the informed consent form). Furthermore, the comparison of the cause of death as reviewed by a clinical expect committee in the NELSON trial with the official death certificates showed a high overall sensitivity and specificity. Therefore, possible biases related to lung cancer death seem relatively small. Moreover, it is suggested that each death case should be reviewed independently by two experts.

SAMENVATTING

Van alle kanker-gerelateerde sterfte is longkanker de voornaamste doodsoorzaak onder mannen en onder vrouwen de een na voornaamste doodsoorzaak. De belangrijkste oorzaak voor longkanker is roken. Ongeveer 85% van de longkankers worden veroorzaakt door roken. Stoppen met roken of nooit beginnen aan is de meest effectieve preventie van longkanker. Ondanks het feit dat de prevalentie van roken is afgenomen, rookt ongeveer een kwart van de Nederlandse bevolking van 18 jaar en ouder. Bovendien wordt tegenwoordig ongeveer de helft van de longkankers vastgesteld in ex-rokers. Klinisch wordt longkanker doorgaans in een laat stadium ontdekt, waarbij vijf jaar na de diagnose slechts 15% nog in leven is. Dit benadrukt de noodzaak voor de vroege opsporing- en behandeling van longkanker, oftewel longkankerscreening. In de Verenigde Staten wordt al gescreend op longkanker. Deze beslissing is deels gebaseerd op de eindresultaten van de grootste studie op het gebied van longkankerscreening - The National Lung Screening Trial (NLST) - en deels op de uitkomsten van een modeleerstudie naar de voor- en nadelen van longkankerscreening welke op verzoek van de United States Preventive Services Task Force (USPSTF) werd uitgevoerd. In Europa wacht men de resultaten af van de grootste Europese studie - de NELSON studie- alvorens een besluit te nemen of longkankerscreening ingevoerd zal worden.

Deel I: optimaliseren van de NELSON screeningsrondes

Het doel van de eerste onderzoeksvraag was om te bepalen welke toegevoegde waarde een vierde, en tevens de laatste, screeningsronde met een interval van 2,5-jaar na de derde screeningsronde heeft (**Hoofdstuk II**). De longkankers gediagnosticeerd in de laatste screeningsronde werden vergeleken met zowel de longkankers gediagnosticeerd in de tweede screeningsronde (met een 1-jaar screeningsinterval) als met de longkankers gediagnosticeerd in de derde screeningsronde (met een 2-jaar screeningsinterval). De longkankers gediagnosticeerd in de vierde screeningsronde zijn statistisch significant in een verder gevorderd stadium vergeleken met de longkankers gediagnosticeerd in de tweede screeningsronde. Hetzelfde werd ook gezien in vergelijking met derde screeningsronde, echter was dat verschil niet statistisch significant. Daarnaast werden er meer intervalkankers gediagnosticeerd na een 2,5-jaar screeningsinterval vergeleken met een 1-jaar en een 2-jaar screeningsinterval. Bovendien werden de meeste intervalkankers in de laatste zes maanden van het 2,5-jaar screeningsinterval gediagnosticeerd.

Vervolgens is er op basis van de eerdere scanuitslagen retrospectief onderzocht of er subgroepen van studiedeelnemers geïdentificeerd konden worden met verschillende risico's op ofwel een niet-negatieve uitslag (dat wil zeggen een twijfelachtige of een positieve scanuitslag) dan wel op longkankerdetectie in de laatste screeningsronde (**Hoofdstuk III**). De kans op een niet-negatieve scan uitslag in de laatste screeningsronde was voor de deelnemers met enkel negatieve scanuitslagen statistisch significant het laagst wanneer dat wordt vergeleken met a) deelnemers met tenminste één twijfelachtige (maar nooit een positieve scanuitslag) uitslag en b) deelnemers met tenminste één positieve scanuitslag. Bovendien was het

longkankerdetectierisico voor de vierde screeningsronde minder dan 1% bij deelnemers met negatieve scanuitslagen in de eerste drie screeningsronden en bij deelnemers met alleen een negatieve uitslag in de derde ronde. Dit suggereert dat deze deelnemers wellicht gedurende een periode van minimaal 2,5 jaar geen screening hoeven te ondergaan. Het gebruik van de screeningshistorie zou mogelijk kunnen leiden tot een betere verdeling van de mogelijke voor- en nadelen voor een groot deel van de te screenen populatie (bijvoorbeeld vermindering van het benodigde aantal scans).

Deel II: tussentijdse vergelijking van de longkankerstadiumverdeling in de NELSON studie

In **Hoofdstuk IV** zijn de stadiumverdeling en behandeling van de in de screen- en controlegroep gediagnosticeerde longkankers onderzocht. Hiervoor zijn de eerste honderd longkankers, zoals gediagnosticeerd in beide groepen, met elkaar vergeleken. Screening middels CT-scan heeft geleid tot de vroegere opsporing van longkanker ten opzichte van geen screening (controle groep). Er werden niet alleen meer longkankers in stadium I gediagnosticeerd, maar ook minder longkankers in stadium IV in de screengroep ten opzichte van de controlegroep. Dit heeft geleid tot meer curatieve behandelingsopties in de screengroep. Concluderend lijkt het, in vergelijking met de NLST, ook mogelijk om middels het gebruik van het NELSON studieprotocol om longkankers in een vroegere stadium te diagnosticeren. Tot zover is het echter onduidelijk of de vroege opsporing van longkankers in de NELSON screengroep-deelnemers ook leidt tot een reductie in longkanker- en totale sterfte.

Deel III: doodsoorzaken onder de NELSON studiedeelnemers

In **Hoofdstuk V** zijn de baselinekenmerken- en sterfte uitkomst van de NELSON controlegroep-deelnemers vergeleken met de 'eligible non-responders' (personen die geschikt waren voor deelname maar uiteindelijk niet hebben deelgenomen of geen toestemming hebben verleend). Zoals verwacht zijn de NELSON studiedeelnemers jonger, gezonder, hoger opgeleid en meer bereid om deel te nemen aan een screeningsprogramma dan de 'eligible non-responders'. Deze verschillen hebben bovendien geleid tot verschil in doodsoorzaak tussen beide groepen: de totale sterfte, onderverdeeld naar sterfte als gevolg van de classificaties kanker, hart- en vaatziekten, longziekte of een andere doodsoorzaak was het hoogst voor de 'eligible non-responders'. Het procentuele verschil was klein en aantal onderzochte personen groot en daarom is het onwaarschijnlijk dat deze verschillen de generaliseerbaarheid van de eindresultaten zal beïnvloeden.

In **Hoofdstuk VI** is er onderzocht of de door een onafhankelijke commissie beoordeelde doodsoorzaken van de NELSON studiedeelnemers overeenkomen met de officiële doodsoorzaken. De onafhankelijke commissie bestond hoofdzakelijk uit twee commissie leden die de geblindeerde medische statussen op een uniforme wijze beoordeelden. Tussen beide commissieleden was de overeenkomst zwak tot middelmatig. De overeenkomst tussen beide commissieleden verbeterde wel naarmate zij meer statussen beoordeelden. Een

vergelijking van de doodsoorzaken zoals geverifieerd door de commissieleden en de officiële doodsoorzaken liet een sensitiviteit en een specificiteit zien van respectievelijk 92.6% en 98.8%. In 12.2% van de casussen werd de doodsoorzaak gewijzigd; waarvan in 46.9% het heeft geleid tot het wijzigen van de doodsoorzaak naar een longkanker-specifieke doodsoorzaak. Een mogelijke vooringenomenheid gerelateerd aan de longkankersterfte lijkt dus relatief klein te zijn in de NELSON studie.

DISCUSSIE

In **Hoofdstuk VII** zijn de hoofdresultaten samengevat en bediscussieerd. Daarnaast zijn achtereenvolgens de methodologische overwegingen, de algemene conclusies van dit proefschrift en de implicaties voor verder onderzoek besproken.

De resultaten in het eerste gedeelte van dit proefschrift toonden aan dat een 2,5-jaar screeningsinterval waarschijnlijk het effect van screening vermindert in vergelijking met een 1-jaar of een 2-jaarscreeningsinterval. Na een 2,5-jaar screeningsinterval werden meer longkankers gediagnosticeerd in een latere stadium en in dit interval werden meer intervalkankers gediagnosticeerd. Daarnaast werden er drie NELSON subgroepen geïdentificeerd met significant verschillende risico's op een niet-negatieve uitslag (dat wil zeggen een twijfelachtige of een positieve scanuitslag) dan wel op longkankerdetectie in de laatste screeningsronde. Hetzelfde werd onderzocht door de NLST, welke in een retrospectieve analyse heeft laten zien dat het inderdaad mogelijk is om subgroepen te identificeren met significant verschillende risico's op longkankerdetectie gebaseerd op hun screeningshistorie. Concluderend, de screeningshistorie zou een behulpzame risico-stratificatietool kunnen zijn om het screeningsprotocol te optimaliseren.

In het tweede gedeelte van dit proefschrift werd aangetoond dat, tot zover, longkanker screening middels CT-scan binnen de NELSON studie heeft geleid tot een substantiële shift in de stadium van de longkankers ten opzichte van geen screening (controle groep). Implementatie van longkankerscreening middels CT-scan zal leiden tot meer longkankerdiagnoses in een vroege stadium en daarbij tot een shift in behandeling waarbij meer curatieve behandelingsmogelijkheden zijn. Alhoewel, het is op dit moment nog onduidelijk of deze gunstige resultaten ook leiden tot een significante longkanker- en totale sterftereductie in de NELSON screengroep-deelnemers.

Zoals verwacht is er een klein 'healthy participant effect' waargenomen in de NELSON studie, welke heeft geleidt tot een verschil in sterfte uitkomst ten opzichte van de 'eligible non-responders'. Verder toonde de vergelijking tussen de door een onafhankelijke commissie beoordeelde doodsoorzaken van de NELSON studiedeelnemers en de officiële doodsoorzaken een hoge sensitiviteit en specificiteit aan. Een mogelijke vooringenomenheid gerelateerd aan de longkankersterfte lijkt dus relatief klein te zijn in de NELSON studie. Tenslotte wordt er aanbevolen dat elk overlijden door minimaal twee commissie leden onafhankelijk van elkaar beoordeeld zou moeten worden.

DANKWOORD

Tot slot het meest gelezen stuk van het proefschrift. Ik wil een groot aantal mensen bedanken die mij hebben geholpen aan de totstandkoming van dit proefschrift. Zonder hun hulp was het niet mogelijk geweest.

Als eerste prof. dr. H.J. de Koning, promotor, de hoofdpersoon tijdens mijn promotieonderzoek. Beste Harry, bedankt dat ik de afgelopen jaren onder uw supervisie en begeleiding hebben mogen werken aan de grootste Europese longkankerscreeningsonderzoek. Net als voor vele anderen bent u voor mij dé expert op het gebied van screening. Ik heb een enorm leuke en vooral een erg leerzame tijd gehad op MGZ, waar ik met veel voldoening op terugkijk. Alles wat ik hierbij heb geleerd neem ik mee in de rest van mijn loopbaan.

Dr. C.M. van der Aalst, co-promotor. Beste Carlijn, samen met Harry was jij één van de belangrijkste personen tijdens mijn promotieonderzoek. Bedankt voor de vrijheid die ik van je kreeg om in mijn eigen tempo te werken aan de NELSON-studie. Bedankt voor alle gezellige gesprekken, zowel over de studie als over het leven. Ik koester al hetgeen ik van je heb mogen leren als onderzoeker maar ook als kersverse moeder. Ontzettend bedankt voor al het geduld dat jij met mij had en om mij door de laatste loodjes heen te lozen.

Prof. dr. J.G.J.V. Aerts, geachte secretaris van de promotiecommissie, als eerste ontzettend bedankt voor het lezen en beoordelen van mijn proefschrift. In het bijzonder wil ik u ook bedanken voor de leuke samenwerking en uw toezegging tot het beoordelen van bijzonder veel medische dossiers voor de "Cause of Death" verificatie van de NELSON studiedeelnemers. Zonder uw inzet zou het behoorlijk lastig zijn geweest om hoofdstuk 6 tot stand te brengen.

Prof. dr. P.E. Postmus en prof. J.W.P.F. Kardaun, geachte leden van de promotiecommissie, ik wil u beiden in het bijzonder bedanken voor het lezen en beoordelen van mijn proefschrift en voor het plaatsnemen in de promotiecommissie.

Prof. M. Oudkerk, bedankt voor de tijd en energie die u in mijn artikelen heeft gestoken. Ik waardeer het dat u zitting wilde nemen in de promotiecommissie.

Prof. dr. I. Beaufort en Prof. dr. S. Siesling, beiden bedankt voor uw tijd en voor het plaatsnemen in de promotiecommissie.

Beste Kevin, inmiddels dr. Ten Haaf, bedankt voor een altijd luisterend oor, je onverbeterlijk positieve instelling, voor alle feedback op mijn werk en vooral bedankt voor al die 'grapjes'. Samen met jou en Carlijn was het altijd gezellig op de afdeling. Ik wens je veel succes toe in je carrière.

Beste Marianne, heel erg bedankt voor al je inspanningen voor de NELSON-studie en alle hulp aan mij. Ik wens je veel succes en liefde toe. Roel en Frank, ik wil jullie beiden

bedanken voor het management van de NELSON-data en voor de begeleiding van alle complexe datakoppelingen. Door de jaren heen ben ik dankzij jullie hulp MS Access steeds meer gaan begrijpen en waarderen. Lieve Nanda, bedankt dat jij mij in het begin wegwijs hebt gemaakt op de afdeling en in alle complexe data van de NELSON-studie. Zonder jouw hulp had ik nooit mijn eerste artikel zo vlot af kunnen hebben. Erik, wat leuk dat jij het longkankerscreeningsteam bent komen versterken. Veel succes met je promotieonderzoek. Lieve Arry, ontzettend bedankt voor het regelen van alle secretariële taken rondom de afspraken met Harry en rondom mijn proefschrift.

Daarnaast wil ik alle NELSON medewerkers, met in het bijzonder prof. dr. H.J.M. Groen, prof. dr. J.W.J. Lammers, prof. dr. P. de Jong, prof. dr. R. Vliegenthart, prof. dr. W. Mali, prof. dr. K. Nackaerts, dr. E.T. Scholten, dr. C. Weenink, dr. F.B.J.M. Thunnissen, dr. M.A. Heuvelmans, J.E. Walter en de lokale datamanagers van alle screeningscentra bedanken.

Dr. M.A. den Bakker wil ik net als prof. J.G.J.V. Aerts bedanken voor het beoordelen van het grote aantal medische dossiers voor de "Cause of Death" verificatie van de NELSON-studiedeelnemers. Hoofdstuk 6 zou zonder jullie hulp een zware uitdaging zijn geweest.

Ook wil ik alle deelnemers van de NELSON studie bedanken voor de essentiële participatie aan het proefbevolkingsonderzoek naar de vroege opsporing van longkanker middels CT-scan.

Het Centraal Bureau van Statistiek, het Nederlands Kankerregister, het Centraal bureau van Genealogie en het Belgisch Kanker Register, bedankt voor de belangrijke bijdrage aan de NELSON-studie. Zonder de benodigde gegevens vanuit deze partijen had ik de publicaties nooit tot stand kunnen brengen.

Beste screensectie, secretariaat en ICT-helpdesk van de afdeling MGZ: allemaal bedankt voor jullie hulp en ondersteuning. Bedankt voor de leuke screensectie- en MGZ-dagen waar elk jaar weer enthousiast naar uitgekeken wordt.

Lieve oud-kamergenoten: Anouk, Epke, Jeroen, Joost, Raquel en Marieke, bedankt voor jullie gezelligheid op de kamer en voor de leuke uitjes. Inmiddels zijn enkele van ons al gepromoveerd. De rest volgt binnenkort. Ik wens jullie heel veel succes met de voorbereiding hiervan.

Daarnaast wil ik de gehele maatschap Interne Geneeskunde/MDL van het Ikazia Ziekenhuis bedanken voor het in mij geïnvesteerde steun en vertrouwen. Ik heb veel mogen leren als co- en arts-assistent bij jullie. Dank daarvoor.

Uiteraard wil ik hierbij ook mijn nieuwe werkomgeving – alle longartsen en arts-assistenten van het Amphia Ziekenhuis – benoemen. Ik ben er maar net, maar het bevalt mij enorm goed. Bedankt voor de prettige en relaxte sfeer op de afdeling, en voor de kans dat ik bij jullie in opleiding ben tot longarts.

Beste Anum en Sheryar, bedankt voor jullie ontzettend grote bijdrage aan het ontwerpen van de kaft van dit proefschrift. Wat zijn jullie een creatief en getalenteerd duo. Ik zal het prachtige en unieke pronkstuk later boven mijn bureau ophangen. Ik wens jullie o.a. veel succes met jullie wetenschappelijke projecten in Pakistan.

Lieve vriendinnen, in het bijzonder Izza, Afza en Noushy, bedankt voor jullie gezelligheid en interesse in mijn promotieonderzoek. Ik hoop dat ik nu meer tijd zal hebben voor dates met onze baby's erbij.

Ten slotte, wil ik mijn familie bedanken voor hun onvoorwaardelijke steun in mij. In het bijzonder wil ik mijn beide broers bedanken die hun zus(je) altijd vol trots hebben gesteund in haar carrière. Mijn zus, zwager en twee neefjes voor hun gezelligheid en gastvrijheid. Mijn mamma wil ik ontzettend bedanken voor haar onvoorwaardelijke liefde en steun aan mij. Mamma, uw bent zo ontzettend ambitieus en talentvol. U bent mijn rolmodel. Al in uw tijd in Pakistan was u erg vooruitstrevend met uw twee masters. Ook nu dat u oma bent, merk ik wat een talent u heeft en dat u nooit stil kunt zitten. Steeds heeft u binnen no time weer een nieuw breistuk klaar voor een van de kleinkinderen. Ik hoop dat u vandaag enorm trots op mij bent. Zonder u was ik nooit zover gekomen. Lieve schoonfamilie, bedankt voor jullie interesse in mijn werk en onderzoek en voor de gezelligheid.

Ahad, mijn lieve echtgenoot en mijn beste vriend. Jouw onvoorwaardelijke liefde, steun en vertrouwen in mij zijn heel belangrijk geweest tijdens mijn promotieonderzoek en tijdens het afronden van dit proefschrift. Bedankt dat jij altijd bereid was om mijn stukken na te lezen op spellings- en grammaticafouten. Inmiddels weet jij al behoorlijk wat van longkankerscreening af. Na deze drukke tijd kijk ik er naar uit om samen weer eens op vakantie te gaan. Bedankt voor alles!

Mijn lieve zoontje Sufian Asher. Wat hou ik ontzettend veel van je. Jouw komst heeft ons leven zo veranderd. Ik hou van je gezwaai en gekus wanneer ik s' ochtend vroeg de deur uit moet. Ik hou van je glimlach en geklap wanneer ik weer eens (te) laat thuis kom van mijn werk. Van deze mooie dag krijg jij waarschijnlijk niet veel mee, maar ik hoop dat jij er later met trots op zal terugkijken. Ik hoop dat ik met mijn successen mooie wegen en mogelijkheden voor jou open heb gelegd.

ABOUT THE AUTHOR

(Assma) Uraujh Yousaf-Khan was born on September 1, 1985 in Rotterdam, the Netherlands. In 2005 she started studying Medicine at the Erasmus University Rotterdam, the Netherlands. She wrote her graduation thesis about the 'Outcomes of BRCA1- and BRCa2-associated versus sporadic breast cancer; the impact on adjuvant systemic therapy' at the department of Internal Oncology at the Daniel den Hoed, Erasmus MC, University Medical Centre Rotterdam. In November 2011, she obtained her medical degree (artsexamen) cum laude. Thereafter, she started working as a resident at the department of Cardiology at Maasstad Hospital in Rotterdam and subsequently as a resident at the department of Internal Medicine at Ikazia Hospital in Rotterdam. From Augustus 2013 to April 2017 she was pointed as junior researcher at the department of Public Health. During this period, she evaluated several aspects of lung cancer screening in the NELSON trial as reviewed in this thesis, under supervision of prof. dr. H.J. de Koning and dr. C.M. van der Aalst.

PHD PORTFOLIO

Summary of PhD training and teaching

Name PhD student: A.U.Yousaf-Khan

Erasmus MC Department: Public Health

PhD period: Aug 2013 – April 2017

Promotor(s):

- Prof. dr. H. J. de Koning
- dr. C.M. van der Aalst

1. PhD training

	Year	Workload (Hours/ECTS)
General courses		
- Biostatistical methods I: basic principles	2013	5.7
- Biostatistical methods II: classical regression models	2013	4.3
- Planning and evaluation of screening	2014	1.4
- Biomedical English Writing and Communication course	2015	3.0
- Research Integrity course	2015	0.3
- BROK ('Basiscursus Regelgeving Klinisch Onderzoek') course	2017	1.4
Seminars and workshops		
- Research meetings / seminars, Erasmus Medical Centre, department of Public Health	2013-2017	2.0
- Workshop "Presenting with theatre skills"	2016	0.2

Presentations

-	Research meeting Pathology, Erasmus Medical Centre: "NELSON longkankerscreeningstudie". Oral presentation	2014	1.0
-	IASCL 2015 Denver (CO), USA: Final screening round of the NELSON study. Oral presentation	2015	1.0
-	Research meeting Public Health, Erasmus MC the NELSON study Oral presentation	2015	1.0

(Inter)national conferences

-	International Cancer Screening Network (ICSN) Meeting	2015	1.0
-	International Association for the Study of Lung Cancer (IASCL)	2015	1.0

2. Teaching

		Year	Workload (Hours/ECTS)
-	Supervising bachelor essays of third year medical students, Erasmus Medical University, Rotterdam	2015	1.1
-	Supervising of Community Projects of the third year medical students, as part of education theme "Arts en volksgezondheid", Erasmus Medical University, Rotterdam	2015	0.6
-	Supervising bachelor essays of third year medical students, Erasmus Medical University, Rotterdam	2016	1.0
-	Supervising of Community Projects of the third year medical students, as part of education theme "Arts en volksgezondheid", Erasmus Medical University, Rotterdam	2016	1.2
-	Supervisor/ tutoring of first year medical students, Erasmus Medical University, Rotterdam	2015-2016	1.4
-	Supervising of Community Projects of the third year medical students, as part of education theme "Arts en volksgezondheid", Erasmus Medical University, Rotterdam	2017	0.6
3. Extracurricular			
-	Introduction in R	2014	1.5 hours
-	Reviewer for the Journal of Thoracic Oncology	2015	3 hours
-	Reviewer for the European Journal of Cancer	2017	3 hours
-	Supervising a medical doctor in data collection from different medical centres and in blinding medical files	2014-2015	-
-	Junior representative for section "Screening"	2015-2016	18 hours
Totaal			**30.1 ECT**

LIST OF PUBLICATIONS

Characteristics of new solid nodules detected in incidence screening rounds of low-dose CT lung cancer screening: the NELSON study. Walter JE, Heuvelmans MA, Bock GH, **Yousaf-Khan U**, Groen HJM, Aalst CMV, Nackaerts K, Ooijen PMAV, Koning HJ, Vliegenthart R, Oudkerk M. Thorax. 2018 Apr 16. pii: thoraxjnl-2017-211376. doi: 10.1136/thoraxjnl-2017-211376. [Epub ahead of print]

Relationship between nodule count and lung cancer probability in baseline CT lung cancer screening: The NELSON study. Heuvelmans MA, Walter JE, Peters RB, Bock GH, **Yousaf-Khan U**, Aalst CMV, Groen HJM, Nackaerts K, Ooijen PMV, Koning HJ, Oudkerk M, Vliegenthart R. Lung Cancer. 2017 Nov;113:45-50. doi: 10.1016/j.lungcan.2017.08.023. Epub 2017 Sep 1.

Uniform and blinded cause of death verification of the NELSON lung cancer screening participants. **A.U.Yousaf-Khan**, C.M. van der Aalst, J.G.J.V. Aerts, M.A. den Bakker, H.J. de Koning. Lung Cancer. 2017 Jul. doi: http://dx.doi.org/10.1016/j.lungcan.2017.07.018. Epub 2017 Jul 22.

Quantification of growth patterns of screen-detected lung cancers: The NELSON study. Heuvelmans MA, Vliegenthart R, de Koning HJ, Groen HJM, van Putten MJAM, **Yousaf-Khan U**, Weenink C, Nackaerts K, de Jong PA, Oudkerk M. Lung Cancer. 2017 Jun;108:48-54. doi: 10.1016/j.lungcan.2017.02.021. Epub 2017 Mar 1.

Risk stratification based on screening history: the NELSON lung cancer screening study. **Yousaf-Khan U**, van der Aalst C, de Jong PA, Heuvelmans M, Scholten E, Walter J, Nackaerts K, Groen H, Vliegenthart R, Ten Haaf K, Oudkerk M, de Koning H. Thorax. 2017 Mar 30. pii: thoraxjnl-2016-209892. doi: 10.1136/thoraxjnl-2016-209892. Epub 2017 Mar 30.

Final screening round of the NELSON lung cancer screening trial: the effect of a 2.5-year screening interval. **Yousaf-Khan U**, van der Aalst C, de Jong PA, Heuvelmans M, Scholten E, Lammers JW, van Ooijen P, Nackaerts K, Weenink C, Groen H, Vliegenthart R, Ten Haaf K, Oudkerk M, de Koning H. Thorax. 2017 Jan;72(1):48-56. doi: 10.1136/thoraxjnl-2016-208655. Epub 2016 Jun 30.

Occurrence and lung cancer probability of new solid nodules at incidence screening with low-dose CT: analysis of data from the randomised, controlled NELSON trial. Walter JE, Heuvelmans MA, de Jong PA, Vliegenthart R, van Ooijen PM, Peters RB, Ten Haaf K, **Yousaf-Khan U**, van der Aalst CM, de Bock GH, Mali W, Groen HJ, de Koning HJ, Oudkerk M. Lancet Oncol. 2016 Jul;17(7):907-16. doi: 10.1016/S1470-2045(16)30069-9. Epub 2016 Jun 6.

Scientific Advances in Lung Cancer 2015. Tsao AS, Scagliotti GV, Bunn PA Jr, Carbone DP, Warren GW, Bai C, de Koning HJ, **Yousaf-Khan AU**, McWilliams A, Tsao MS, Adusumilli PS, Rami-Porta R, Asamura H, Van Schil PE, Darling GE, Ramalingam SS, Gomez DR,

Rosenzweig KE, Zimmermann S, Peters S, Ignatius Ou SH, Reungwetwattana T, Jänne PA, Mok TS, Wakelee HA, Pirker R, Mazières J, Brahmer JR, Zhou Y, Herbst RS, Papadimitrakopoulou VA, Redman MW, Wynes MW, Gandara DR, Kelly RJ, Hirsch FR, Pass HI. J Thorac Oncol. 2016 May;11(5):613-38. doi: 10.1016/j.jtho.2016.03.012. Epub 2016 Mar 22. Review.

Baseline Characteristics and Mortality Outcomes of Control Group Participants and Eligible Non-Responders in the NELSON Lung Cancer Screening Study. **Yousaf-Khan U**, Horeweg N, van der Aalst C, Ten Haaf K, Oudkerk M, de Koning H. J Thorac Oncol. 2015 May;10(5):747-53. doi: 10.1097/JTO.0000000000000488.

Detection of lung cancer through low-dose CT screening (NELSON): a prespecified analysis of screening test performance and interval cancers. Horeweg N, Scholten ET, de Jong PA, van der Aalst CM, Weenink C, Lammers JW, Nackaerts K, Vliegenthart R, ten Haaf K, **Yousaf-Khan UA**, Heuvelmans MA, Thunnissen E, Oudkerk M, Mali W, de Koning HJ. Lancet Oncol. 2014 Nov;15(12):1342-50. doi: 10.1016/S1470-2045(14)70387-0. Epub 2014 Oct 1.