

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

Uraujh Yousaf-Khan,¹ Carlijn van der Aalst,^{1,2} Pim A de Jong,² Marjolein Heuvelmans,^{3,4} Ernst Scholten,^{2,5} Jan-Willem Lammers,⁶ Peter van Ooijen,^{3,4} Kristiaan Nackaerts,⁷ Carla Weenink,⁸ Harry Groen,⁹ Rozemarijn Vliegenthart,^{3,4} Kevin ten Haaf,¹ Matthijs Oudkerk,^{3,4} Harry de Koning¹

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/thoraxjnl-2016-208655>).

For numbered affiliations see end of article.

Correspondence to

Uraujh Yousaf-Khan, Erasmus Medical Center, University Medical Center Rotterdam, Department of Public Health, Room Na 22-09, P.O. Box 2040, Rotterdam 3000 CA, The Netherlands; a.yousaf@erasmusmc.nl

Received 21 March 2016

Revised 9 May 2016

Accepted 2 June 2016

Published Online First

30 June 2016

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 $\geq 25\%$ compared with no screening after 10 years of follow-up. The screening arm (n=7915) 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.

Trial registration number ISRCTN63545820.

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),^{2,3} the US Preventive Services Task Force recommends annual lung cancer screening with CT. People eligible for screening are aged 55 years through

Key messages

What is the key question?

- What is the additional value of a fourth screening round with an interval of 2.5 years after the previous three screening rounds.

What is the bottom line?

- An interval of 2.5 years leads to a higher interval cancer rate and a higher proportion of advanced stage disease in the final fourth round compared with the previous screening rounds with a 1-year or 2-year screening interval.

Why read on?

- The Dutch-Belgian Lung Cancer Screening trial (NELSON) is the largest European randomised lung cancer screening trial which is sufficiently powered and it provides a unique opportunity to investigate the effect of variable screening intervals (1 year, 2 years and 2.5 years intervals) in one screening group.

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=1190 and n=1186).^{6,7}

The Dutch-Belgian Lung Cancer Screening trial (NELSON) is the largest European randomised lung cancer screening trial, which was designed to investigate whether LDCT screening reduces lung cancer mortality by $\geq 25\%$ 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



► <http://dx.doi.org/10.1136/thoraxjnl-2016-209011>



CrossMark

To cite: Yousaf-Khan U, van der Aalst C, de Jong PA, et al. *Thorax* 2017;**72**:48–56.



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).

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 $\geq 25\%$ 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 7915 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: $<50 \text{ mm}^3$ was negative, $50\text{--}500 \text{ mm}^3$ was indeterminate and $>500 \text{ mm}^3$ 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 $\geq 25\%$, 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

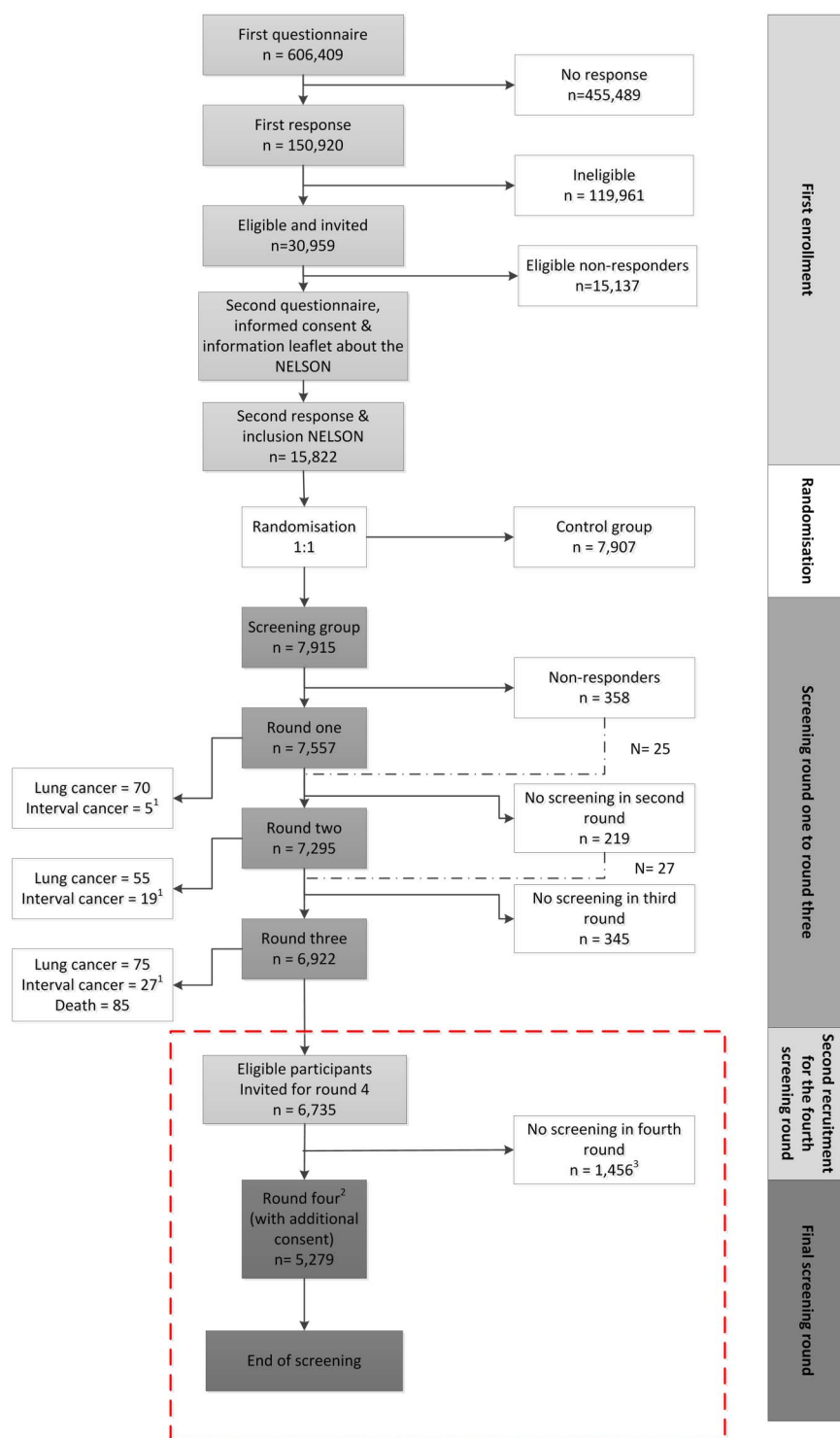
A regular round scan is the first CT examination performed for a specific participant in one of the predefined screening rounds. Follow-up scans are repeat scans which were performed in between screening rounds if a participant had an indeterminate result in one of the four regular scans. The result 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.

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. ¹Interval cancer data of only Dutch screening group participants. ²Only participants who gave their additional consent were screened in the last screening round. ³Reasons 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 1353 participants didn't respond. ⁴Eventually, 5,279 participants provided additional consent and were screened in the final screening round.



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 χ^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.

RESULTS

The fourth screening round

The participation rates in the first three rounds were: 7557 (95.5%) in round 1, 7925 (92.2%) in round 2 and 6922 (87.5%) in round 3. All eligible participants from the third round were invited to participate in the final round; in total 6735 of the 6922 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% (5437/6735) responded positively, and 97.1% (5279/5437) 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: 5380 scans were performed of which 98.1% (5279/5380) were regular round scans, and 1.9% (101/5380) were follow-up scans performed to assess the VDT of intermediate nodules.

In total 5336 nodules were detected on 2507 of 5380 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–IIa) compared with small-cell lung cancers which were diagnosed in stage IIIa or stage IV

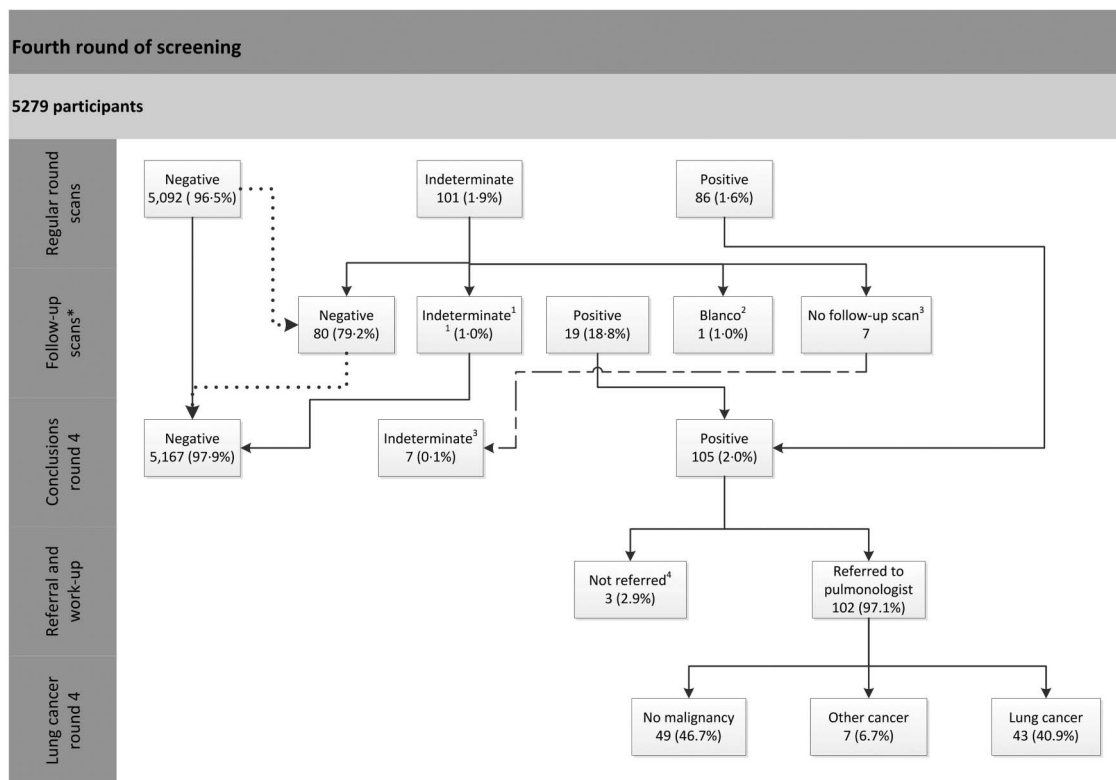


Figure 2 Screening results of the fourth screening round. *Follow-up scans are scans performed after an indeterminate screening result. Dotted line: four participants with a negative regular scan result received accidentally a follow-up scan. Their screening results remained negative. ¹In the follow-up scan one participant received an indeterminate result, so the participant received a second follow-up scan: the final scan was negative. ²From one participant the scan data were lost, the participant received a new follow-up scan. The final screening result was negative. ³Seven 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. ⁴Three 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 workup 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	Per cent
Total scans performed in round 4	5380	100.0
Scans with nodules	2507	46.6
Scans without nodules	2873	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	3057	57.3
NODCAT III***		
Indeterminate	1875	35.2
NODCAT IV****		
Potentially malignant	149	2.8
Nodule growth category†		
GROWCAT§ A†	2348	98.1
GROWCAT§ B‡	14	0.6
GROWCAT§ C‡‡	29	1.3
Type		
Solid	5012	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§	3109	58.3
Left lung	2227	41.7

*NODCAT I: Nodule with benign characteristics such as benign calcification patterns or fat depositions.

**NODCAT II: Solid nodules with volume <50 mm³; Pleural-based solid nodules with minimum diameter <5 mm; Non-solid component part solid nodule with average diameter <8 mm; Solid component part solid nodule with volume <50 mm³; Non-solid nodules with average diameter <8 mm.

***NODCAT III: Solid nodules with volume 50–500 mm³; Pleural-based solid nodules with minimum diameter 5–10 mm; Non-solid component part solid nodule with average diameter ≥8 mm; Solid component part solid nodule with volume 50–500 mm³; Non-solid nodules with average diameter ≥8 mm.

****NODCAT IV: Pleural-based solid nodules with minimum diameter >10 mm; Solid component part solid nodule with volume >500 mm³.

†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.

(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).

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.

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 2 Clinical features of screening-detected lung cancers in the fourth screening round

	Cancer stage							Total histology n, %
	Ia n, %	Ib n, %	IIa n, %	IIb n, %	IIIa n, %	IIIb n, %	IV 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: 47.8% (11/23) of all adenocarcinomas were detected in stage Ia.

†For example 23 (50.0) should be read as: 50% (23/46) of all screening-detected lung cancers in this round were adenocarcinomas.

‡Of which one was a pTisNOMx 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: 47.8% (22/46) of the cancers detected in this round were stage Ia lung cancers.

BAC, bronchoalveolar carcinoma; NSCLC NOS, non-small cell lung carcinoma not otherwise specified.

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	Per cent	Cumulative %		n	Per cent	Cumulative %		n	Per cent	Cumulative %		n	Per cent	Cumulative %
				0.37				0.02				0.10			
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.

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	Per cent		n	Per cent		n	Per cent		n	Per cent
			0.01			0.001			0.06		
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	2	2.7		3	5.2		6	7.8		4	8.7
SCLC	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 rounds. Fisher's exact test was used.

†p Value: difference in histology between the second and last screening rounds. Fisher's exact test was used.

‡p Value: difference in histology between the third and last screening rounds. Fisher's exact test was used.

§Others in round 1: 5 large cell carcinomas, 1 mixed NSCLC/SCLC, 4 carcinoids, 1 other, 3 no diagnosis was possible, 2 adenosquamous and 3 large-cell neuroendocrine 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 neuroendocrine 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; NSCLC, non-small cell lung carcinoma; SCLC: small-cell carcinoma.

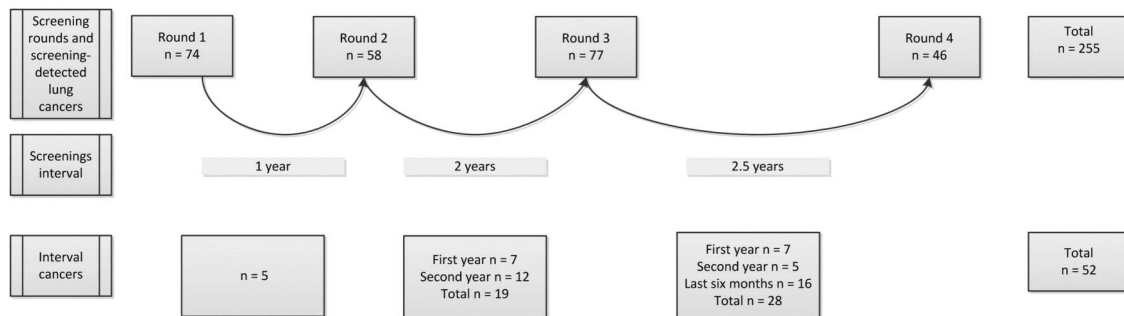
Table 5 Screening test performance across the four screening rounds

	R1*	95% CI	R2*	95% CI	R3*	95% CI	R4	95% CI
Lung cancer detection rate, %	0.9	0.7 to 1.2	0.8	0.6 to 1.0	1.1	0.8 to 1.3	0.8	0.6 to 1.1
Positive predicted value, %	35.5	28.4 to 42.1	42.0	34.4 to 49.6	45.5	37.6 to 53.5	41.0	31.6 to 50.5
False-positive (FP) rate after positive screening, %	64.5	57.9 to 71.6	58.0	50.4 to 65.6	54.5	46.7 to 62.4	59.0	49.5 to 68.4
Ratio TP/FP [†]	0.69	–	0.72	–	0.83	–	0.69	–
Overall FP rate [†]	–	–	–	–	–	–	1.2 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.**Table 6** Baseline characteristics of participants

Characteristics	Participants with interval cancer in the 2.5-year interval after round 3 (n=27)*	Participants screened in round 4 (n=5279)	Participants with screening-detected lung cancer in round 4 (n=43)	p Value
Male n, (%)	23 (85.2)	4437 (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)	2878 (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 cancers 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 1 participant with an interval lung cancer (2nd year round 3) were diagnosed with synchronous double tumours.

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% (105/5380) 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). 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 randomised 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.

Author affiliations

¹Department of Public Health, Erasmus Medical Center, University Medical Center Rotterdam, Rotterdam, The Netherlands

²Department of Radiology, University Medical Center Utrecht, Utrecht, The Netherlands

³Department of Radiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

⁴University of Groningen, University Medical Center Groningen, Center for Medical Imaging—North East Netherlands, Groningen, The Netherlands

⁵Department of Radiology, Kennemer Gasthuis, Haarlem, The Netherlands

⁶Department of Pulmonary Medicine, University Medical Center Utrecht, Utrecht, The Netherlands

⁷Department of Pulmonary Medicine, University Hospital Gasthuisberg, Leuven, Belgium

⁸Department of Pulmonary Medicine, Kennemer Gasthuis, Haarlem, The Netherlands

⁹Department of Pulmonary Diseases, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

Acknowledgements The authors thank the system controllers R Faber and FJP Santegeerts, 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.

Contributors Involvement in the conception, hypothesis delineation and design of the study: UY-K, CvdA, MO and HdK. Acquisition of the data or the analysis and interpretation of such information: UY-K, CvdA, PADJ, ES, PVO, RV, MO and HdK. Writing the article or substantial involvement in its revision prior to submission: UY-K, CvdA, PADJ, MH, ES, J-WL, PVO, KN, CW, HG, RV, KtH, MO and HdK. UY-K had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Funding The NELSON trial is supported by: The NELSON trial is presently supported by "The Netherlands Organisation for Health Research and Development" (ZonMw). Siemens Germany provided four digital workstations and LungCARE for the performance of 3D-measurements of the nodules. Roche diagnostics provided a grant for the performance of proteomics research.

Competing interests KN reports grants from Belgian Foundation against Cancer, grants from Flemish League against Cancer, during the conduct of the study. HG reports others from MSD, Eli Lilly, Merck, Pfizer, Roche, outside the submitted work. HdK reports grants from Roche Diagnostics, other from Siemens Germany, during the conduct of the study; personal fees from MD Anderson/Health sciences, outside the submitted work. UY-K, KtH and CvdA report grants from Roche Diagnostics, and other from Siemens Germany, during the conduct of the study.

Patient consent Obtained.

Ethics approval The NELSON trial was approved by the Dutch Minister of Health and the ethics board at each participating centre. The NELSON trial is registered at <http://www.trialregister.nl> under number: ISRCTN63545820.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

- 1 Siegel R, Ma J, Zou Z, *et al.* Cancer statistics, 2014. *CA Cancer J Clin* 2014;64:9–29.
- 2 Aberle DR, Berg CD, Black WC, *et al.*, National Lung Screening Trial Research Team. The National Lung Screening Trial: overview and study design. *Radiology* 2011;258:243–53.

- 3 Aberle DR, Adams AM, Berg CD, *et al.*. National Lung Screening Trial Research Team. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med* 2011;365:395–409.
- 4 Moyer VA, U.S. Preventive Services Task Force. Screening for lung cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2014;160:330–8.
- 5 de Koning H, Meza R, Plevritis S, *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;160:311–20.
- 6 Pastorino U, Rossi M, Rosato V, *et al.* Annual or biennial CT screening versus observation in heavy smokers: 5-year results of the MILD trial. *Eur J Cancer Prev* 2012;21:308–15.
- 7 Sverzellati N, Silva M, Calareso G, *et al.* Low-dose computed tomography for lung cancer screening: comparison of performance between annual and biennial screen. *Eur Radiol*. Published Online First: 11 Feb 2016. doi:10.1007/s00330-016-4228-3
- 8 van Iersel CA, de Koning HJ, Draisma G, *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;120:868–74.
- 9 van Klaveren RJ, Oudkerk M, Prokop M, *et al.* Management of lung nodules detected by volume CT scanning. *N Engl J Med* 2009;361:2221–9.
- 10 Xu DM, Gietema H, de Koning H, *et al.* Nodule management protocol of the NELSON randomised lung cancer screening trial. *Lung Cancer* 2006;54:177–84.
- 11 Horeweg N, van der Aalst CM, Thunnissen E, *et al.* Characteristics of lung cancers detected by computer tomography screening in the randomized NELSON trial. *Am J Respir Crit Care Med* 2013;187:848–54.
- 12 Horeweg N, van der Aalst CM, Vliegenthart R, *et al.* Volumetric computed tomography screening for lung cancer: three rounds of the NELSON trial. *Eur Respir J* 2013;42:1659–67.
- 13 Horeweg N, Nackaerts K, Oudkerk M, *et al.* Low-dose computed tomography screening for lung cancer: results of the first screening round. *J Comp Eff Resp* 2013;2:433–6.
- 14 Horeweg N, Scholten ET, de Jong PA, *et al.* Detection of lung cancer through low-dose CT screening (NELSON): a prespecified analysis of screening test performance and interval cancers. *Lancet Oncol* 2014;15:1342–50.
- 15 Xie X, Willemink MJ, Zhao Y, *et al.* Inter- and intrascanner variability of pulmonary nodule volumetry on low-dose 64-row CT: an anthropomorphic phantom study. *Br J Radiol* 2013;86:20130160.
- 16 Xie X, Zhao Y, Snijder RA, *et al.* Sensitivity and accuracy of volumetry of pulmonary nodules on low-dose 16- and 64-row multi-detector CT: an anthropomorphic phantom study. *Eur Radiol* 2013;23:139–47.
- 17 Wang Y, van Klaveren RJ, van der Zaag-Loonen HJ, *et al.* Effect of nodule characteristics on variability of semiautomated volume measurements in pulmonary nodules detected in a lung cancer screening program. *Radiology* 2008;248:625–31.
- 18 Ru Zhao Y, Xie X, de Koning H, *et al.* NELSON lung cancer screening study. *Cancer Imaging* 2011;11(Spec No A):S79–84.
- 19 van der Aalst C, van Klaveren R, de Koning HJ, *et al.* Does participation to screening unintentionally influence lifestyle behaviour and thus lifestyle-related morbidity? *Best Pract Res Clin Gastroenterol* 2010;24:465–78.
- 20 Saghir Z, Dirksen A, Ashraf H, *et al.* CT screening for lung cancer brings forward early disease. The randomised Danish Lung Cancer Screening Trial: status after five annual screening rounds with low-dose CT. *Thorax* 2012;67:296–301.
- 21 Aberle DR, DeMello S, Berg CD, *et al.* Results of the two incidence screenings in The National Lung Screening Trial. *N Engl J Med* 2013;369:920–31.
- 22 Wille MM, Dirksen A, Ashraf H, *et al.* Results of the randomized Danish lung cancer screening trial with focus on high-risk profiling. *Am J Respir Crit Care Med* 2016;193:542–51.