Screening for lung cancer by imaging
Oudkerk, Matthijs; Heuvelmans, Marjolein

Published in:
Jbr-Btr

DOI:
10.5334/jbr-btr.240

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2013

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Lung cancer is a major health problem with no improvement in survival over the last decades. At time of diagnosis, lung cancer is often already in advanced stage, with 5-year survival of no more than 15% (1). Currently, several lung cancer screening trials investigating whether early detection of lung cancer in high-risk individuals will eventually reduce lung cancer mortality are ongoing (2-7). To date, the National Lung Screening Trial (NLST) is the only randomized controlled trial in which a significant lung cancer mortality reduction was found (2).

The Dutch-Belgian lung cancer screening trial (Dutch acronym: NELSON study) was launched in September 2003. The NELSON study is an ongoing multicentre randomized controlled multi-detector low-dose computed tomography lung cancer screening trial. The primary object is to investigate whether chest CT screening in year 1, 2, 4 and 6.5 will decrease lung cancer mortality by at least 25% in high-risk (ex-)smokers between 50 and 75 years of age compared to a control group receiving no screening. Secondary end points of the study include estimation of the cost-effectiveness of the screening programme, assessment of the optimal screening interval (1, 2 or 2.5 years), and assessment of the impact on quality of life. In addition, multiple side studies are ongoing.

One of the major challenges in lung cancer screening is the high false-positive rate, causing patient anxiety, cost and morbidity associated with unnecessary diagnostic procedures for benign nodules. The NELSON trial is the first large lung cancer screening trial in which the nodule management protocol is based on nodule volume, instead of nodule diameter, and nodule growth, in terms of volume doubling time (VDT) of existing nodules. The final results will indicate whether a volumetry- and VDT based CT protocol is more efficient in terms of detection rate, morbidity, mortality, recall rate, and cost-effectiveness, compared to other approaches.

Methods

Participants

The NELSON multi-centre trial was approved by the Dutch Minister of Health and the ethics board at each participating centre. All participants provided written informed consent. Participants were recruited based on a questionnaire about health, smoking, cancer history, and other lifestyle and health factors. Included were current or former heavy smokers, with a history of > 15 cigarettes daily for > 25 years or > 10 cigarettes daily for > 30 years and between 50-75 years of age. Exclusion criteria were a moderate or bad self reported health, inability to climb two flights of stairs, body weight > 120 kg, lung cancer < 5 years ago or still under treatment, current or past renal cancer, melanoma or breast cancer, and chest CT less than 1 year ago (8). In total, 15,822 subjects were included. 7,557 were categorized in the screen group, receiving low-dose chest CTs. Participants in the control group received no screening.

Participants in the screen group underwent chest CT, and depending on the screening round, pulmonary function testing and blood sampling on the same day. After each CT examination, participants completed a quality of life questionnaire.

Data acquisition

The participants randomized to the screen group were invited to one of the four screening sites (University Hospital Groningen, University Hospital Utrecht and Kennemer Gasthuis Haarlem in the Netherlands and University Hospital Gasthuisberg Leuven in Belgium). All low-dose chest CT scans were performed by using 16-detector helical CT scanners Sensation-16, Siemens Medical Systems or, at the screening site in Utrecht, Mx8000 IDT or Brilliance 16P, Philips Medical Systems). Scanning of the entire chest was performed in a caudo-cranial direction, without the use of contrast agents. Depending on body weight (< 50 kg, 50-80 kg, and > 80 kg), the kVp settings were 80-90, 120 and 140 kVp respectively. This corresponds to an effective radiation dose < 1.6 mSv. Data sets of the lung were reconstructed at 1.0-mm slice thickness, with 0.7-mm reconstruction increment. Scans were performed in inspiration after appropriate instruction of the participants, to minimize breathing artefacts.

Volumetric measurements and image reading

Digital workstations (Leonardo, Siemens Medical Solutions) were used for nodule volumetric analysis. This system detected automatically whether a nodule, marked by a radiologist, was new or had been present previously. After a nodule was marked, a program for semi-automated volume measurements (LungCare, version Somaris/5 VB 10A-W, Siemens Medical Solutions) automatically defined the volume of interest around the nodule. An observer could manually modify the segmentation by increasing or decreasing the volume, if necessary (9).

Data generated by the LungCare software were uploaded into the NELSON Management System, which automatically detected whether a nodule was new or present on previous scans. The percentage volume change and VDT of previously detected nodules were calculated...
automatically by the system. For each evaluable nodule, the surface characteristics, location, distance to the pleura, and aspect of the nodule (i.e., solid, non-solid or partial solid) were entered in the NELSON Management System by a radiologist.

All CT images were independently read by first and second readers (double reading) as part of the NELSON protocol (9). The first reading was performed by a reader with experience in reading chest CTs varying from none to more than 20 years; the second reading was performed by two readers, each with 6 years of experience. The second readers were unaware of the conclusions of the first readers. In case of discrepancy, the final decision was made by a third reader (10).

**Screening strategy**

At baseline, a test was considered positive if any non-calcified nodule was larger than 500 mm³ (>9.8-mm diameter). The result was indeterminate if the volume of the largest solid nodule or the solid component of a partially-solid nodule was 50-500 mm³ (4.6-9.8-mm diameter). In case of smaller nodules, the screening was negative (9).

Indeterminate nodules underwent a 3-month follow-up CT to assess for growth. Growth was defined as volume increase of at least 25%. For growing nodules, the final result was based on their VDT. If a growing lesion had a VDT < 400 days, the final result was positive. Otherwise the baseline result was negative and the participant was invited for the regular second-round examination in year 2.

At second-round screening, there were two possibilities: either a nodule was new, and the result was based on nodule size, or a nodule was pre-existing. New indeterminate nodules underwent a 6-weeks follow-up CT. For pre-existing nodules, the second round result was based on their VDT immediately. If both new and existing nodules were present, the nodule with the largest volume or fastest growth determined the result. Again, a VDT < 400 days resulted in a positive screen result. A nodule with VDT > 600 was classified as negative. A VDT of 400-600 days comprised an indeterminate result; then a follow-up CT was made 1 year later. Then, if the VDT was less than 400 days, the final result was positive (Fig. 1), otherwise negative. All participants with a negative second-round result were invited to undergo the third round examination 2 years after the second round (9).

The protocols of the third and fourth-round examinations were comparable to the protocol of the second-round, except for the fact that the fourth-round examination was planned 2.5 years after the third-round examination. The screening program ended after a positive or negative fourth round result, or, in case of an indeterminate fourth-round result, after a positive or negative follow-up CT (Fig. 2).

Test positives were referred to a pulmonologist for workup. Workup, staging, and treatment were standardized to (inter-) national guidelines (9, 11). Nodules were classified as benign or malignant based on histological examination. Also, nodules could be classified benign based on stable or decreasing size two years after first detection (12, 13). If lung cancer was diagnosed, the participant was treated and left the screening trial; otherwise the regular next-round CT was scheduled.

**Results**

**NELSON screen results**

The results of the first and second screening round were published in...
In 2009 the screen group had a nodule with volume 1.6% of the subjects in the group received 3 annual rounds of low-dose CT screening. The control group received 3 annual rounds of chest X-ray screening (2).

In the NLST screening rounds, the rate of positive tests, defined as greatest nodule diameter of 4 mm or larger, was 24%. No less than 96.4% comprised false positive results. Volume-based nodule management has been suggested to be more accurate than diameter measurements (14, 15), potentially leading to lower false-positive rates. Therefore, the NELSON trial was the first lung cancer screening trial which based screening interpretation on nodule volumetry and growth in terms of volume doubling time instead of diameter measurements. This strategy yielded a rather low rate of positive screening tests (2.6% in the baseline screening; 1.8% in the second-round screening), while the number of missed lung cancers was low.

**Additional results of the NELSON study**

Valuable knowledge about interobserver variability and the optimal image reading protocol of semi-automated volumetry was obtained in the NELSON trial. Gietema et al. found that interobserver correlation was very high (r = 0.99) in small-to-intermediate size (15-500 mm³) lung nodules (10). It was also found that variability on volume measurements is related to nodule size, morphology and location (16). In a further study (17), a difference in repeatability among three reconstruction settings was found. It was shown that volume measurement of pulmonary nodules obtained at 1 mm section thickness combined with a soft kernel was most repeatable. Therefore it was concluded that in case of serial CT studies, consistent reconstruction parameters are essential. Furthermore, compared to single reading, no statistically significant benefit for consensus double reading at baseline screening for lung cancer with the use of a nodule management strategy based solely on semi-automated volumetry was found (18). Therefore, in the fourth screening round image reading was performed by only one reader.

At last, the performance of computer aided detection (CAD) was compared to double reading. The false-positive rate was 3.7 per CT for CAD and 0.5 per CT for readers. Excluding small nodules (<50 mm³), the false-positive rate for CAD decreased to 1.8. The sensitivity of nodule detection by readers for nodules with need of further evaluation could have increased by 18.6% (from 78.1% to 96.7%) if CAD had also been used. However, only one lung cancer missed by readers was detected by CAD (19).

Three studies focussed on the work-up of pulmonary nodules. In the first (20), it was shown that conventional white-light bronchoscopy should not be routinely recommend for patients with a positive test result in a lung cancer screening trial. The overall sensitivity was 13.5% and the negative predictive value was 47.6%. In the second study (21), the role of a preoperative positron emission tomography after a conclusive or inconclusive nonsurgical workup was evaluated. It was concluded that a preoperative PET scan in participants with an inconclusive nonsurgical workup is not recommended because of the very low negative predictive value. The third study (22) investigated the complication rate in participants of the screen arm of the NELSON lung cancer screening trial who underwent surgical resection. They found that mortality rates after surgical procedures are lower in the NELSON lung cancer screening trial than those in the non-screening series. The role of complications is within the same range as in the non-screening series.

A number of studies focussed on the characteristics of lung nodules associated with cancer risk. In solid nodules larger than 50 mm³, especially size, and to a lesser extent irregular shape and margin, were found to increase the likelihood of malignancy (13). Although baseline lung

---

**Fig. 2.** — Screening programme

---

The New England Journal of Medicine in 2009 (7). In the baseline screening round, 1.6% of the subjects in the screen group had a nodule with volume >5000 mm³. 19.2% had at least one indeterminate nodule, for which a three-month follow-up CT was performed. In this follow-up CT, growth was demonstrated in only 5.3% of participants with indeterminate nodules. In total, 196/7557 participants tested positive (2.6%). In 70/196 participants, malignancy was confirmed; the lung cancer detection rate was 0.9%. Sensitivity of the baseline round screening was 94.6%, the negative predictive value was 99.9%. In the second round screening, the rate of positive tests, defined as greatest nodule diameter of 4 mm or larger, was 24%. No less than 96.4% comprised false positive results. Volume-based nodule management has been suggested to be more accurate than diameter measurements (14, 15), potentially leading to lower false-positive rates. Therefore, the NELSON trial was the first lung cancer screening trial which based screening interpretation on nodule volumetry and growth in terms of volume doubling time instead of diameter measurements. This strategy yielded a rather low rate of positive screening tests (2.6% in the baseline screening; 1.8% in the second-round screening), while the number of missed lung cancers was low.

**Additional results of the NELSON study**

Valuable knowledge about interobserver variability and the optimal image reading protocol of semi-automated volumetry was obtained in the NELSON trial. Gietema et al. found that interobserver correlation was very high (r = 0.99) in small-to-intermediate size (15-500 mm³) lung nodules (10). It was also found that variability on volume measurements is related to nodule size, morphology and location (16). In a further study (17), a difference in repeatability among three reconstruction settings was found. It was shown that volume measurement of pulmonary nodules obtained at 1 mm section thickness combined with a soft kernel was most repeatable. Therefore it was concluded that in case of serial CT studies, consistent reconstruction parameters are essential. Furthermore, compared to single reading, no statistically significant benefit for consensus double reading at baseline screening for lung cancer with the use of a nodule management strategy based solely on semi-automated volumetry was found (18). Therefore, in the fourth screening round image reading was performed by only one reader.

At last, the performance of computer aided detection (CAD) was compared to double reading. The false-positive rate was 3.7 per CT for CAD and 0.5 per CT for readers. Excluding small nodules (<50 mm³), the false-positive rate for CAD decreased to 1.8. The sensitivity of nodule detection by readers for nodules with need of further evaluation could have increased by 18.6% (from 78.1% to 96.7%) if CAD had also been used. However, only one lung cancer missed by readers was detected by CAD (19).

Three studies focussed on the work-up of pulmonary nodules. In the first (20), it was shown that conventional white-light bronchoscopy should not be routinely recommended for patients with a positive test result in a lung cancer screening trial. The overall sensitivity was 13.5% and the negative predictive value was 47.6%. In the second study (21), the role of a preoperative positron emission tomography after a conclusive or inconclusive nonsurgical workup was evaluated. It was concluded that a preoperative PET scan in participants with an inconclusive nonsurgical workup is not recommended because of the very low negative predictive value. The third study (22) investigated the complication rate in participants of the screen arm of the NELSON lung cancer screening trial who underwent surgical resection. They found that mortality rates after surgical procedures are lower in the NELSON lung cancer screening trial than those in the non-screening series. The role of complications is within the same range as in the non-screening series.

A number of studies focussed on the characteristics of lung nodules associated with cancer risk. In solid nodules larger than 50 mm³, especially size, and to a lesser extent irregular shape and margin, were found to increase the likelihood of malignancy (13). Although baseline lung...
nodule CT density was not predictive of malignancy, an increase in CT density on follow-up CTs in intermediate-sized nodules suggested lung cancer (23).

Cancers in intermediate-sized (50-500 mm³) fast-growing solid nodules, diagnosed at 3-month or 1-year follow-up CT after baseline, were found to be non-spherical and purely intraparenchymal, without attachment to the pleura, vessels or fissures (24). Perifissural nodules, accounting for about 20% of all lung nodules found in lung cancer screening, can show growth rates in the range of malignant nodules. However, none of the perifissural nodules turned out to be malignant after 5.5 years of follow-up. Therefore, recognition of these nodules can reduce unnecessary workup (25).

Conclusion

The first results of the NELSON study show the value of 3D-based lung nodule management for CT lung cancer screening, with very high negative predictive values found in the first and second screening rounds of the NELSON study population is ongoing and the mortality results are pending, but the unique methodological features of this randomized trial have already yielded important insights that complement the information gained from NLST.

References