

Risk assessment in relation to the detection of small pulmonary nodules

John K. Field¹, Michael W. Marcus¹, Matthijs Oudkerk²

¹Roy Castle Lung Cancer Research Programme, Department of Molecular and Clinical Cancer Medicine, Institute of Translational Medicine, The University of Liverpool, Liverpool, UK; ²University of Groningen, University Medical Center Groningen, Center for Medical Imaging EB 45, 9700RB Groningen, the Nederland

Contributions: (I) Conception and design: JK Field, M Oudkerk; (II) Administrative support: JK Field, MW Marcus; (III) Provision of study materials or patients: JK Field, M Oudkerk; (IV) Collection and assembly of data: JK Field, M Oudkerk; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: John K. Field, PhD, FRCP. Roy Castle Lung Cancer Research Programme, Department of Molecular and Clinical Cancer Medicine, Institute of Translational Medicine, The University of Liverpool, Liverpool, UK. Email: J.K.Field@liverpool.ac.uk; Matthijs Oudkerk, MD, PhD, FACA, EBCR. University of Groningen, University Medical Center Groningen, Center for Medical Imaging EB 45, 9700RB Groningen, the Nederland. Email: m.oudkerk@umcg.nl.

Abstract: The National Lung Cancer Screening trial (NLST) demonstrated that individuals assigned to the LDCT screening arm had a 20% lower mortality than those who were assigned to the conventional chest radiography. The NLST was thoroughly analyzed by the US Preventive Task Force on CT Screening and they recommended that lung cancer screening should be implemented. A number of other countries have also recommended implementation, whilst others are awaiting the outcome of the NELSON Trial. However, recommendations for the management of CT screen detected nodules have only recently had any clarity. The management of CT detected nodules in the NLST was based on the identification and reporting of 4 mm diameter nodules found on the CT screens but there was no NLST radiology protocol in place for the management of nodules. The use of volumetric analysis is not routinely used in the USA and there is still a reliance on utilising the CT nodule diameter as the management parameter. The first pulmonary risk model was developed by the Canadians, utilising data sets from the Pan-Canadian Early detection of Lung cancer (PanCan) and validated in the chemoprevention trial dataset at the British Columbian Agency. This Canadian model, known as the Brock Model, is currently available and has been integrated into the British Thoracic Society guidelines on the management of pulmonary nodules. The American College of Radiology setup a Lung Cancer Screening Committee subgroup on Lung-RADS, to standardize lung cancer screening CT reporting and provide management recommendations. However, it has been recommended that the Lung-RADS system should be revised as the system as it has never been studied in a prospective fashion. The NELSON trial introduced a third screening test, the “indeterminate” screening test result, this was done with the aim to reduce the false-positives CT screening results and also utilized by the UKLS trial successfully. On comparing the radiological CT screen volumetric and diameter based protocols in the NELSON trial, the sensitivity and negative predictive value appeared to be comparable, however a higher specificity and positive predictive value was found for the volume-based protocols, thus confirming the advantage of utilising the volumetric approach over diameter. The British Thoracic Society (BTS) has undertaken an in-depth piece of work developing guidelines on the management of pulmonary nodules, utilising the wealth of data published by the NELSON team and support the use of volumetric analysis for the management of pulmonary nodules.

Keywords: Lung cancer; risk prediction; pulmonary nodules; management

Submitted Jan 14, 2017. Accepted for publication Jan 17, 2017.

doi: 10.21037/tlcr.2017.02.05

View this article at: <http://dx.doi.org/10.21037/tlcr.2017.02.05>

The use of low dose CT (LDCT) for early lung cancer detection in high risk individuals has progressed from the first publication by Henschke *et al.* (1), through to the National Lung Cancer Screening trial (NLST) in 2011 (2), to the current data from the NELSON trial on the management of small pulmonary nodules. The NLST demonstrated that individuals assigned to the LDCT screening arm had a 20% lower mortality than those who were assigned to the conventional chest radiography. The current status of lung cancer screening trials has been extensively reviewed over the past three years, demonstrating the enormous strides in the management of lung cancer screening (3-6).

Clearly the stage has been set in the USA for the implementation of lung cancer screening based on the NLST trial publication and also on the recommendation from the US Preventive Services Task Forces (USPSTF) (7) on lung cancer screening, resulting in the agreed funding from March 2016 by the Center for Medicare (8) and Medicaid (CMS). The USPSTF recommended annual screening for lung cancer in the 55–80 age group who have a 30-pack-year smoking history and were either current smokers or have quit within the last 15 years. The independent review set up by the USPSTF modelled screening policies and investigated the long-term harms and benefits of lung cancer screening. The USPSTF have indicated that the parameters for selection should be review in time together with the management of these patients.

We currently await the publication of the NELSON

trial, which will provide valuable information on mortality and cost effectiveness, from the only fully powered European trial. However, all of the main CT screening trials have consistently demonstrated that early Stage disease is one of the core findings, with 81% from International Early Lung Cancer Detection Program (IELCAP), 63% from NLST, 73% from NELSON and 67% from the (United Kingdom Lung Screening (UKLS) trial (*Table 1*), compared to the expected ~15%. It also note that a number of pilot European CT screening trials have provided an in-depth insight into the management of CT detected nodules.

The management of CT detected nodules in the NLST was based on the identification and reporting of 4 mm diameter nodules found on the CT screens but there was no NLST radiology protocol in place for the management of nodules. Clearly, the early work undertaken by IELCAP initiated the debate on utilising volumetric measurements for the management of small CT detected nodules. This work has been further developed by the NELSON group and latterly validated by the UKLS trial.

The use of volumetric analysis is not routinely used in the USA and there is still a reliance on utilising the CT nodule diameter as the management parameter. The Canadian Pulmonary Risk model was developed utilising datasets from the Pan-Canadian Early detection of Lung cancer (PanCan) and validated in the chemoprevention trial dataset at the British Columbian Agency (BCCA) (9).

Characterisation of nodules is well described within the PanCan risk model publication included a range of

Table 1 Early stage cancers identified in lung cancer RCT trials

Trial	Participants in screening arm	Screening rounds	No. published CT detected lung cancers	Stage IA & IB lung cancers (%)
NLST	26,722	3	649	400 (61.6)
NELSON	7,915	4	209	148 (70.8)
DLST	2,052	5	69	47 (68.1)
ITALUNG	1,613	4	22	11 (50.0)
DANTE	1,276	4	58	41 (70.7)
MILD	1,190	10	20	18 (62.1)
	1,186	5	22	14 (70.0)
LUI SI	2,029	4	22	18 (81.8)
UKLS	1,994	1	42	28 (67.8)
Total	54,977	1–10	1,120	725 (64.0)

NLST, National Lung Cancer Screening trial.

imaging parameters including spiculation, which was found to be a major predictor in the PanCan dataset, however, was not confirmed within the BCCA, as this data was not collected. The authors went on to develop parsimonious and full models with and without nodule spiculation. The model's discrimination i.e. a measure of how well such model can separate diseased from non-diseased individuals is most often measured using the area under the receiver characteristic (ROC) curve or c-statistic (10). Halligan *et al.* has identified problems with ROC and argued that it depends on the method used for curve fitting and does not account for prevalence or different misclassification costs arising from false-negative and false-positive diagnoses (11). Other methods and metrics of the performance of prediction models, such as the net benefit, have been proposed based on the change in sensitivity and specificity at clinical relevant thresholds (12). A major strength of this model is that it does not solely rely on ROC because comparison of the models with and without spiculation showed no significant differences in AUC but the net re-classification between the two models did suggest that spiculation could improve prediction. Net benefit incorporates estimates of prevalence and misclassification costs, and it is clinically interpretable since it reflects changes in correct and incorrect diagnoses when a new diagnostic test is introduced (11,12). The take home message was that if a threshold of at least 5% risk of lung cancer is used in the parsimonious model including spiculation, the sensitivity, specificity, positive predictive value and negative predictive value were: 71.4%, 95.5%, 18.4% and 99.6%. Thus, the model developed by McWilliams *et al.* can be used to accurately estimate the probability that lung nodules detected on baseline screening with low-dose CT scans are malignant. This model showed good accuracy for determining likelihood of malignancy in nodules detected on CT scans (13). However, in patients undergoing (fluorodeoxyglucose positron emission tomography-computed tomography) FDG PET-CT for nodule evaluation, the highest accuracy was seen in the Herder and co-workers risk model (14).

Lung CT screening Reporting and Data System (Lung-RADS_{ver1}) was published in 2014 (15). The American College of Radiology setup a Lung Cancer Screening Committee subgroup on Lung-RADS, in order to have a quality assurance tool to standardize lung cancer screening CT reporting and also provide management recommendations. The rationale behind this initiative is the hope that it would assist in lung cancer screening CT nodule scan interpretations. However, when Lung-RADS

performance was compared to the NLST screening trial data, certain issues arose, even though NLST summary data was used to construct the Lung-RADS scores (16). The comparative performance indicated that Lung-RADS substantially reduced the false positive result rate and the sensitivity level decreased. Recently it has been recommended by Mehta *et al.* that the Lung-RADS system needs to be revised and faulted the system on the basis that it has never been studied in a prospective study.

Li *et al.* have recently analysed the size and growth of pulmonary nodules, as a consequence of 'rounding up' methodology used in Lung-RADS (17). The example given is if a nodule with an average diameter of 5.5 mm is reported as 6 mm diameter since 6 mm diameter is the current threshold for a positive result, further workup would be recommended for this nodule. Thus, rounding up to the nearest whole number increases the frequency of positive results which require further work-up before the next scheduled screening round. The authors also indicated another possible confusion, as to whether the length or the width is rounded up, which is not indicated in the Lung-RADS criteria. The authors concluded that with the move towards the utilisation of computer aided techniques, rounding up will be used less often, furthermore, the trend towards volumetric assessment of nodules, will result in a much more precise methodology.

The NELSON trial introduced a third screening test, the indeterminate screening test result, this was done with the aim to reduce the false-positives CT screening results (18). The importance of this decision is seen in the low percentage of false positives found in the NELSON trial. Especially, when one looks at the impact of the false positive screening test, with potential unnecessary work-up and invasive procedures and the possibility of overtreatment and the extra anxiety for the patients.

In the UKLS (19), a very clear definition was made for false positive tests as those requiring further diagnostic investigation more immediately than a repeat annual screen, but who subsequently did not have lung cancer. The proportion of false positive tests was provided in two ways, which allows an appreciation, in a patient-centered approach, of the variable impact on the subject in a trial or the patient in a programme. A "false positive" that mandates referral to the lung cancer multidisciplinary team (MDT) clinic will usually be associated with significant psychological distress, and additional invasive investigations with, in some cases, definitive treatment. An individual with a false positive as defined above is more likely to suffer harm

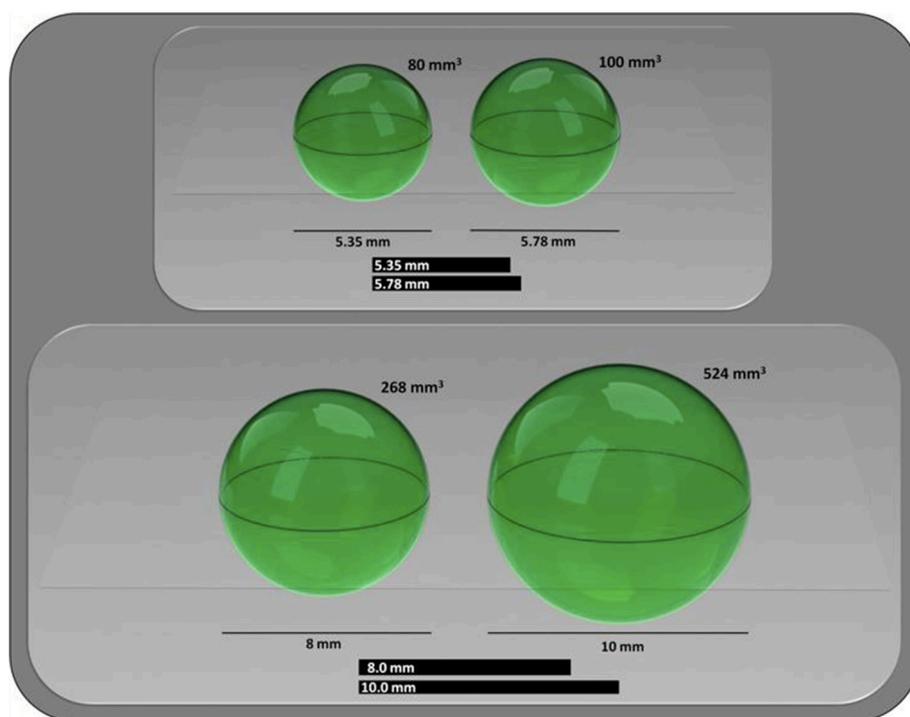


Figure 1 Comparison of the diameter and volume of CT screen detected nodules. (A) A volume growth of 26%, defined as growth by NELSON criteria, is hardly appreciable by diameter measurement (8% diameter increase which is NO growth by current criteria); (B) a 25% diameter increase i.e., threshold for the current growth definition reflects almost a doubling in volume (95%). It reflects the insensitivity for growth of diameter measurement. Reproduce from reference (3).

than one defined in a different way; that is, those subjects who are recalled solely for further CT imaging to clarify the nature of a nodule. The latter is best termed “Interval Imaging Rate” and may, in screening programmes, merely mean continuing in the programme rather than referral to the MDT. For this reason, all category three lesions in the UKLS trial without cancer (or called indeterminate nodules) were reported separately as false positives warranting interval imaging (19).

In the UKLS, the false positive rate was 3.6% whilst the interval imaging rate was 23.2% amongst participants referred to MDT clinic. The NELSON trial reported their false positive rate in 2013 as 3.6% (20). Both the UKLS and the NELSON utilised the indeterminate screening result whenever the participant received a repeat test within a period of three months, which was analysed by utilising volumetric analysis. A 25% increase in volume was considered as ‘nodule growth’ and the patient was then referred to the MDT for conventional clinical work-up. The advantage of utilising volumetric analysis is diagrammatically demonstrated in *Figure 1*. On comparing the radiological

CT screen volumetric and diameter based protocols in the NELSON trial, the sensitivity and negative predictive value appeared to be comparable, however a higher specificity and positive predictive value was found for the volume-based protocols (21) thus confirming the advantage of utilising the volumetric approach over diameter.

The data discussed so far in this article relates to baseline data with nodule follow-up, however, the trial data which is relevant for routine screening are on new and incidental nodules comes from the extensive work undertaken by the NELSON team. NELSON calculated the risk of developing lung cancer based on the volume, volume based diameter in a large dataset of screened participants found to have non-calcified nodules and developed a probability table (*Figure 2*). It’s of note that the probability was not significantly different between the NELSON participants with nodules <100 mm³ compared to those with no CT detected nodules in the trial (0.6% *vs.* 0.4%). However, individuals with 100–300 mm³ nodal volume had a higher probability of developing lung cancer (2.4%) and were considered indeterminate with intermediate risk; whilst the participants with nodules greater

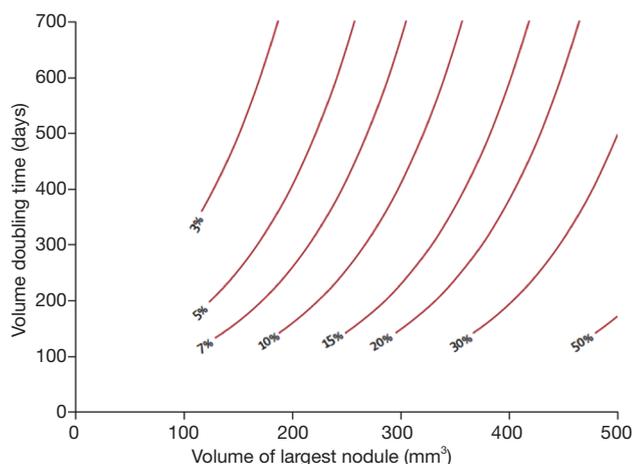


Figure 2 Contour plot of the effect of the combined effect of nodule volume and volume doubling time on 2-year lung cancer probability. The risk isolines represent the percentage of NELSON participants that will be diagnosed with lung cancer within 2 years according to the volume of their largest nodule and volume doubling time of the fastest growing nodule in the 50–500 mm³ range. Reproduce from reference (21).

than 300 mm³ had a significantly greater risk compared to no nodules (16.9%) and thus had a very high probability of developing lung cancer (21).

A very important message was provided on examining the NELSON volume doubling time data; the 2-year probability of developing lung cancer in patients with nodules measuring 50–100 mm³ (or 4–5 mm diameter) was extremely low and did not significantly differ from patients with no CT scan detected nodules. This observation questions whether these individuals require yearly CT scans in a long term screening program and takes into account the harm and benefits for regular screening in such individuals; i.e., radiation exposure, psychological distress and cost effectiveness.

New pulmonary nodules at incident screens are now recognised as a clinical issue which has been analysed by Walters *et al.* (22). NELSON registered 1,222 new nodules in 787 participants. Fifty lung cancers were found, representing 4% of all new solid nodules and 34 (68%) lung cancers were diagnosed at stage I. They reported that the new nodules with <27, 27–206, 206 mm³ were classified as low (0.5%), intermediate (3.1%) and high risk (16.9%) probability of developing lung cancer. The NELSON data showed that new solid nodules are detected at each

screening round in 5–7% of patients and have a significant probability of being malignant, even if they are of small size. These findings will have an impact on the way we develop our future screening guidelines.

The British Thoracic Society (BTS) has undertaken an in-depth piece of work developing guidelines on the management of pulmonary nodules (23). This work has been based on extensive review of the literature and the utilisation of recent publication from a number of lung cancer CT screening trials and in-depth analysis of data. A Guideline Development Group (GDG) was assembled utilising new research evidence, they have provided four management algorithms and they have included two malignancy prediction calculators (already discussed in this article) (Figure 3). Furthermore, volumetry has been recommended by BTS as the preferred measurement method of CT detected nodules and they also provided recommendations for the management of nodules with extended volume doubling times.

The BTS guidelines provide recommendation on the use of further imaging, and the use of PET-CT information which can be incorporated into pulmonary risk models, as well as advice on biopsy and the threshold for treatment without histological confirmation. Finally, BTS provided advice on the information which should be given to patients on the management of pulmonary nodules.

Clearly, the field of pulmonary nodule management in CT screening continues to advance and with the recent publication on the risk of malignancy in new nodules which has highlighted the need to continuously refine the nodule management algorithms and that the new nodule risk data should be taken into account (24).

Lung cancer screening is now a reality in the USA, covered by the Center for Medicare and Medicaid, however, Europe and the rest of the world have not yet implemented national lung cancer CT screening programmes at the time of writing this article, as they await the publication of the NELSON trial, with its mortality and cost effectiveness data. There will be a range of challenges when each country starts to implement lung cancer screening programmes in Europe, which have already been identified (25) but we also need to ensure that the appropriate protocolled pulmonary nodule management pathways such as the BTS recommendations are agreed and put in place, in order that we achieve the greatest clinical impact from future lung cancer screening programmes.

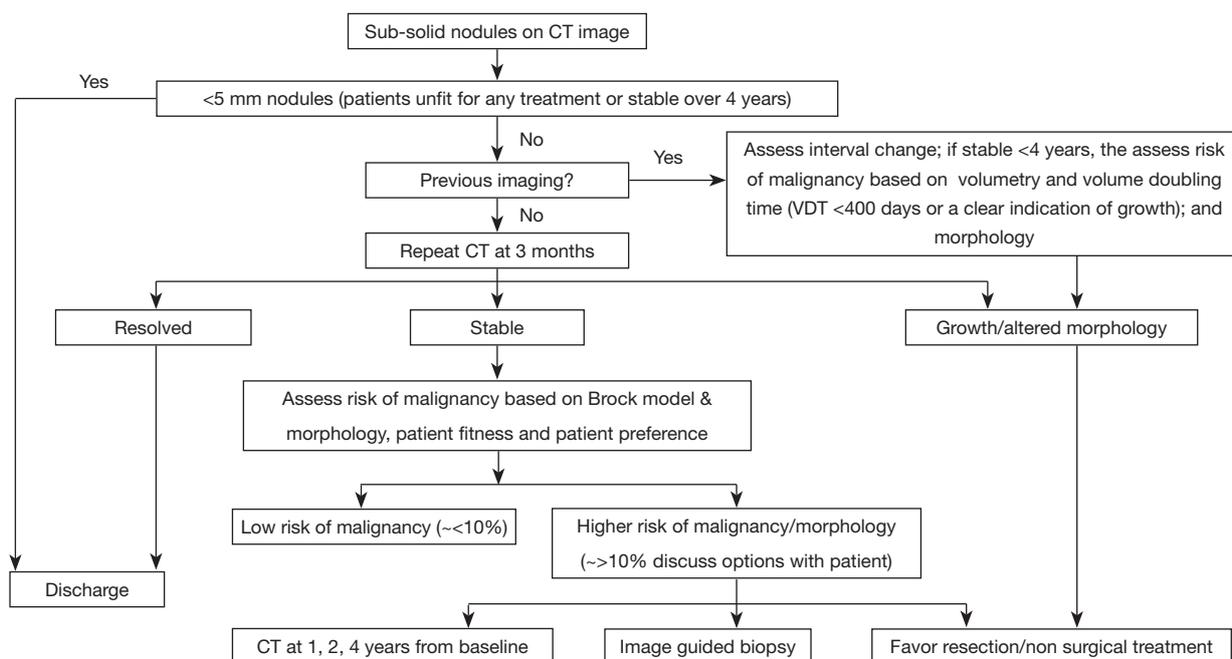


Figure 3 Management of pulmonary nodules. Redrawn from reference (23).

Acknowledgements

None.

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

References

- Henschke CI, McCauley DI, Yankelevitz DF, et al. Early Lung Cancer Action Project: overall design and findings from baseline screening. *Lancet* 1999;354:99-105.
- National Lung Screening Trial Research T, Aberle DR, Adams AM, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med* 2011;365:395-409.
- Field JK, Oudkerk M, Pedersen JH, et al. Prospects for population screening and diagnosis of lung cancer. *Lancet* 2013;382:732-41.
- Field JK, Hansell DM, Duffy SW, et al. CT screening for lung cancer: countdown to implementation. *Lancet Oncol* 2013;14:e591-600.
- Field JK. Perspective: The screening imperative. *Nature* 2014;513:S7.
- Field JK, Devaraj A, Duffy SW, et al. CT screening for lung cancer: Is the evidence strong enough? *Lung Cancer* 2016;91:29-35.
- USPSTF Lung Cancer Screening [Internet]. 2015 [cited 25-11-2016]. Available online: <http://www.uspreventiveservicestaskforce.org/Page/Document/UpdateSummaryFinal/lung-cancer-screening>
- Decision Memo for Screening for Lung Cancer with Low Dose Computed Tomography (LDCT) (CAG-00439N) [Internet]. Available online: <https://www.cms.gov/medicare-coverage-database/details/nca-decision-memo.aspx?NCAId=274>. Cited 25-11-2016.
- McWilliams A, Tammemagi MC, Mayo JR, et al. Probability of cancer in pulmonary nodules detected on first screening CT. *N Engl J Med* 2013;369:910-9.
- Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation* 2007;115:928-35.
- Halligan S, Altman DG, Mallett S. Disadvantages of using the area under the receiver operating characteristic curve to assess imaging tests: a discussion and proposal for an alternative approach. *Eur Radiol* 2015;25:932-9.
- Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*

- 2010;21:128-38.
13. Al-Ameri A, Malhotra P, Thygesen H, et al. Risk of malignancy in pulmonary nodules: A validation study of four prediction models. *Lung Cancer* 2015;89:27-30.
 14. Herder GJ, van Tinteren H, Golding RP, et al. Clinical prediction model to characterize pulmonary nodules: validation and added value of 18F-fluorodeoxyglucose positron emission tomography. *Chest* 2005;128:2490-6.
 15. Fintelmann FJ, Bernheim A, Digumarthy SR, et al. The 10 Pillars of Lung Cancer Screening: Rationale and Logistics of a Lung Cancer Screening Program. *Radiographics* 2015;35:1893-908.
 16. Pinsky PF, Gierada DS, Black W, et al. Performance of Lung-RADS in the National Lung Screening Trial: A Retrospective Assessment. *Ann Intern Med* 2015;162:485-91.
 17. Li K, Yip R, Avila R, et al. Size and Growth Assessment of Pulmonary Nodules: Consequences of the Rounding. *J Thorac Oncol* 2016. [Epub ahead of print].
 18. 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.
 19. Field JK, Duffy SW, Baldwin DR, et al. The UK Lung Cancer Screening Trial: a pilot randomised controlled trial of low-dose computed tomography screening for the early detection of lung cancer. *Health Technol Assess* 2016;20:1-146.
 20. 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.
 21. Horeweg N, van Rosmalen J, Heuvelmans MA, 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;15:1332-41.
 22. Walter JE, Heuvelmans MA, de Jong PA, 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. *Lancet Oncol* 2016;17:907-16.
 23. Callister ME, Baldwin DR, Akram AR, et al. British Thoracic Society guidelines for the investigation and management of pulmonary nodules. *Thorax* 2015;70 Suppl 2:ii1-ii54.
 24. Baldwin DR, Devaraj A. Lung cancer risk in new pulmonary nodules: implications for CT screening and nodule management. *Lancet Oncol* 2016;17:849-50.
 25. Field JK, Duffy SW, Devaraj A, et al. Implementation planning for lung cancer screening: five major challenges. *Lancet Respir Med* 2016;4:685-7.

Cite this article as: Field JK, Marcus MW, Oudkerk M. Risk assessment in relation to the detection of small pulmonary nodules. *Transl Lung Cancer Res* 2017;6(1):35-41. doi: 10.21037/tlcr.2017.02.05