Lessons on managing pulmonary nodules from NELSON: we've come a long way

C Horst[1], A Nair[1], S M Janes[1]

The results of the Dutch-Belgian low-dose CT (LDCT) screening trial (NELSON), have been eagerly awaited since the publication of the US National Lung Screening Trial in 2011 demonstrated annual low dose CT screening of the chest led to a 20% decrease in lung cancer mortality compared to CXR screening(1). While the United States Preventative Service Task Force approved lung cancer screening in 2014(2), Europe and the rest of the globe have been paralysed by fear of implementation costs and the feasibility of introducing national LDCT screening programmes. The consensus from healthcare payers outside of the US has been that we should wait for the results of the NELSON trial, which while smaller in size, would give us the confidence of a second randomised controlled trial and proof of effect in a population outside of the US healthcare system. Indeed, a recent Health Technology Assessment of low-dose CT screening in the UK specifically named the NELSON trial as an important source of future information, and explicitly stated that their results were required in order to make a decision about its efficacy and cost-effectiveness(3). It was with this background of enthusiastic hope and perhaps, dare we say it, health care payer fear, that the NELSON trial preliminary results were released at the World Conference on Lung Cancer in Toronto in October.

Despite NELSON being smaller in size than NLST and having a preponderance of male participants, the results were clear: LDCT screening compared to no screening leads to a statistically significant lung cancer mortality reduction of 26% for men; and numbers hint that the benefit could be even greater in women (between 40-60%). With this new data, the UK, and indeed the world outside of the US, now needs to cast aside concerns over efficacy, as well as procrastination over implementation, and concentrate more on how we can successfully deliver a cost-efficient screening programme reaching those at highest risk.

One of several key areas of work when running a LDCT programme is dealing efficiently and accurately with the management of pulmonary nodules, where the need to diagnose lung cancer early is compromised by the danger of over-investigating benign nodules and driving up both financial cost and potential participant harms. New pulmonary nodules appearing in individuals with previously negative CT screening rounds are particularly vexing. On the one hand, these nodules are most likely transient and due to some (likely subclinical) infection or inflammation; on the other, a fast-growing lung cancer could easily present as a new nodule on an annual interval scan. At an individual level,

[1] Lungs for Living Research Centre, UCL Respiratory, University College London, London, UK

there should be no dilemma as to what to do: large nodules require immediate investigation, while smaller pulmonary nodules require further surveillance CT to establish their growth rate. However, when extrapolated to screening populations numbering in their tens or even hundreds of thousands, the problem becomes more complex, as bringing everyone with a new nodule back for surveillance CT could incur alarming costs for a population screening programme. The question then becomes: are current thresholds of size and growth applicable to new nodules as well as old, or are more aggressive cut-offs required for such nodules, to distinguish between benign and malignant lesions?

Many studies have attempted to do just that by evaluating different nodule characteristics, including morphology, nodule location, size (both diameter and volume), growth rate (including volume doubling time) and other, more esoteric measures (e.g. radiomics and computer learning) (4–11). All of these investigations are pieces of the pulmonary nodule jigsaw puzzle which, we hope, will fit together to form a complete picture of how to risk stratify and manage these lesions. It is certain that the evidence around nodule behaviour will continue to be collected, and the conclusions drawn from this evidence will continue to be refined, particularly where computer-aided tools are concerned. But given the new impetus to establish national screening programmes, we will need to use the best evidence we currently have to support these programmes at their inception.

The NELSON team has led the way in providing evidence for nodule management; in this journal and others, they have delineated the behaviour of pulmonary nodules detected both at baseline, and at incidence rounds of screening(12–16). They rightly point out that screening is not a single event but a continuous process; serial scanning, as part of a screening programme, will show the evolution (and possible devolution) of nodules at set intervals, and that nodules identified at baseline and subsequent rounds may require different management strategies. In two previous papers on the subject of new solid nodules, they have shown that: (1) the size (volume) threshold for new solid nodules found at incidence rounds which require monitoring should be lowered (to approximately 30mm^3), as should the threshold for immediate investigation (to 200mm^3, from 300mm^3), as such nodules are more likely to be malignant(12); (2) new solid nodules in a central or right upper lobe location, and solid nodules which were <15 mm^3 in retrospective reviews of prior CTs (in other words, sub-threshold but not truly new nodules) are significant predictors of lung cancer(16).

In this issue of Thorax, the NELSON group present an epilogue to these prior investigations (Walter, J, et al. Persisting new nodules found at incidence rounds of the NELSON lung cancer screening study, Thorax 2019, epub ahead of print), focussing again on risk stratification. This manuscript reinforces many of the key messages from the earlier publications. It is reassuring that the proportion of new solid nodules that are malignant is low and consistent with their previous

publications: 25 non-resolving new solid nodules were cancer, corresponding to 7.0% of the 356 participants with non-resolving new solid nodules. Also, the threshold for immediately referring a new solid nodule at first discovery- 200mm^3 or larger- is reinforced. Finally, for those of us worried about tiny new nodules, it is salutary that small nodules (<50mm^3) visible in retrospect had no higher chance of being cancer than new nodules below this size threshold: 97% [224/232] of the nodules visible in retrospect were <50mm3 (ie less than 5mm in spherical nodule diameter) at initial detection and the lung cancer probability (1.3% [3/224], CI 0.3-4.0%) was similar compared to new solid nodules <50mm^3 that were not visible in retrospect (1.5% [6/394], CI 0.6-3.4%, (P=0.855).

What, then, is new in this latest offering? Firstly, the current analysis finally answers a a key clinical question: how many new nodules simply resolve? The answer is the majority: 55% of incident nodules resolved, in 47% of participants found to have them (the discrepancy between these numbers is accounted for by some people having multiple new nodules). This brings nuance to the debate around one of the criticisms levelled against screening, namely the 'high false positive rate'. Indeed, NLST quoted a rate of 23.3%(1), and numbers coming out some screening centres in the US are almost twice that(17). But the fact NELSON has shown that most new nodules that appear at incident scans subsequently disappear again lends weight to the argument that we should not regard them automatically as 'positive' findings, but instead treat them as findings requiring surveillance, to be assessed via imaging, but not invasively investigated. This has long been the position of LDCT screening advocates in the UK and Europe, where the term 'interval imaging rate' is used to indicate nodules requiring surveillance, as opposed to automatically categorising nodules of a certain size as 'positive findings'.(5)

Secondly, the NELSON investigators assessed the diagnostic utility of stratifying nodules by volume doubling time (VDT) using two approaches: the three-category VDTs of their original management protocol (VDT <400 days, 400-600 days, and >600 days) and an optimised cut-off derived from their data, of 590 days. Further, they investigate the utility of integrating a VDT ≤590 days with the previously established high-risk volume threshold of ≥200mm^3. Interestingly, this combination outperformed volume alone but was not significantly better than VDT alone. Considering a nodule high-risk for malignancy when it had a VDT ≤590 days or volume ≥200mm^3 augmented sensitivity to 100% but at the expense of specificity (84%) for lung cancer, as opposed to 92% sensitivity and 87% specificity using VDT ≤590 days alone. Tellingly, the original analysis of new solid nodules did not demonstrate the discriminatory value of VDT(12), underscoring the value of reappraising data.

The authors argue that these results provide justification for urgent referral to a pulmonologist when a non-resolving new solid nodule has a VDT ≤590 days. This latter

recommendation deserves special attention, because 22/25 (88%) of the new solid nodules that were lung cancer had a VDT<400 days while only 1/25(4%) lung cancers had a VDT of ≥400 days but <600 days (Table S5 in the publication); this latter lung cancer was diagnosed in a non-resolving nodule on a subsequent LDCT performed more than 120 days later. As the upper limit of the interquartile range of VDTs in lung cancer nodules of this subgroup was 362 days (Table 1 in the publication), it is safe to say that the growth rate of this single lung cancer with a VDT of 400-600 days was an outlier. Even if a nodule measuring 30-199mm^3 (the indeterminate range for new solid nodules on incident screening rounds) had a VDT of 400-590 days at a 3 month follow-up CT, and was managed conservatively with a subsequent CT follow-up at a year 1, such a nodule would at most have grown to 375mm^3- an increase from 7 to 9mm for a 199mm^3 spherical nodule. Thus, it is hard to imagine what a pulmonologist or multidisciplinary team conference would do with such a nodule if referred at 3 months, other than relegate it to further follow-up. The recommendation to more aggressively pursue nodules with VDT ≤590 days rather than the established cut-off of ≤400 days, then, albeit based on receiver operator characteristic and accuracy metrics analyses, is perhaps a case of the statistical tail being allowed to wag the dog a little too strongly.

This fact notwithstanding, the compendium of NELSON analyses does elegantly illustrate three important and complementary lessons. First, nodule management is far from a discovered country, and continuous re-appraisal of data can still lead to new insights; for this, the NELSON investigators deserve our continued gratitude. Second, even with new insights, refinement of existing volume- and growth-based strategies, be it on incidence or baseline screening rounds, will still not translate into a panacea that achieves both 100% sensitivity and specificity in population-based LDCT lung cancer screening. Finally, the ability of such continued re-appraisal of the same data from the same dataset to refine management strategies will, at some point, inevitably plateau.

The most headline-worthy results of the NELSON trial are still awaited in print form, but that should not detract from the important task of assembling the best evidence for how to take screening forward. Indeed, in a post-NELSON world, where the mortality benefit of LDCT screening has been confirmed, it is even more imperative that the screening community reaches a consensus on how best to manage nodules, both at baseline and at incidence rounds. This is not least because a screening programme lives or dies on its ability to reduce harms, minimise false positives and identify cancers early; only by putting together the pulmonary nodule jigsaw puzzle, piece by piece, will we be able to argue that the exciting lung cancer mortality reduction is truly worth the effort.

1. Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, Fagerstrom RM. Reduced lung-cancer

mortality with low-dose computed tomographic screening. N Engl J Med. 2011;365.

2. Moyer VA. Screening for Lung Cancer: U.S. Preventative Services Task Force Recommendation Statement. Ann Intern Med. 2014;160(5):330–8.

3. Snowsill T, Yang H, Griffin E, Long L, Varley-Campbell J, Coelho H, et al. Low-dose computed tomography for lung cancer screening in high-risk populations: a systematic review and economic evaluation. Health Technol Assess (Rockv). 2018;22(69):1–276.

4. McWilliams A, Tammemagi MC, Mayo JR, Roberts H, Liu G, Soghrati K, et al. Probability of Cancer in Pulmonary Nodules Detected on First Screening CT. N Engl J Med. 2013;369(10):910–9.

5. Field JK, Duffy SW, Baldwin DR, Whynes DK, Devaraj A, Brain KE, et al. UK Lung Cancer RCT Pilot Screening Trial: Baseline findings from the screening arm provide evidence for the potential implementation of lung cancer screening. Thorax. 2016;71(2):161–70.

6. Lee SH, Lee SM, Goo JM, Kim KG, Kim YJ, Park CM. Usefulness of texture analysis in differentiating transient from persistent part-solid nodules(PSNs): A retrospective study. PLoS One. 2014;9(1).

7. Wilson R, Devaraj A. Radiomics of pulmonary nodules and lung cancer. Transl Lung Cancer Res. 2017;6(1):86–91.

8. Nibali A, He Z, Wollersheim D. Pulmonary nodule classification with deep residual networks. Int J Comput Assist Radiol Surg. 2017;12(10):1799–808.

9. Armato SG, Drukker K, Li F, Hadjiiski L, Tourassi GD, Engelmann RM, et al. LUNGx Challenge for computerized lung nodule classification. J Med Imaging. 2016;3(4):044506.

10. Talwar A, Willaime JMY, Pickup LC, Enescu M, Boukerroui D, Hickes W, et al. Pulmonary nodules: Assessing the imaging biomarkers of malignancy in a “coffee-break.” Eur J Radiol. 2018;101(February):82–6.

11. Talwar A, Gleeson F, Rahman NM, Pickup L, Gooding M, Kadir T. Validation Of 4 Models To Estimate The Probability Of Malignancy In Patients With Sub-Centimeter Pulmonary Nodules. Am J Crit Care Med. 2015;191:A6117.

12. Walter JE, Heuvelmans MA, de Jong PA, Vliegenthart R, van Ooijen PMA, 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. Lancet Oncol.

2016;2045(16):1–10.

13. Heuvelmans MA, Walter JE, Oudkerk M. Management of baseline and new sub-solid nodules in CT lung cancer screening. Expert Rev Respir Med. 2018;12(1):1–3.

14. 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.

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

16. Walter JE, Heuvelmans MA, Bock GH De, Yousaf-khan U, Groen HJM, Aalst CM Van Der, et al. Characteristics of new solid nodules detected in incidence screening rounds of low-dose CT lung cancer screening : the NELSON study. Thorax. 2018; 73: 741–7.

17. Caverly TJ, Fagerlin A, Wiener RS, Slatore CG, Tanner NT, Yun S, et al. Comparison of Observed Harms and Expected Mortality Benefit for Persons in the Veterans Health Affairs Lung Cancer Screening Demonstration Project. JAMA Intern Med. 2018;1–3.