

Keywords: lung cancer; screening; estimates of effect; modelling

Translation of research results to simple estimates of the likely effect of a lung cancer screening programme in the United Kingdom

S W Duffy^{*1}, J K Field², P C Allgood¹ and A Seigneurin³

¹Centre for Cancer Prevention, Wolfson Institute of Preventive Medicine, Queen Mary University of London, Charterhouse Square, London EC1M 6BQ, UK; ²Roy Castle Lung Cancer Research Programme, The University of Liverpool Cancer Research Centre, Roy Castle Building, 200 London Road, Liverpool L3 9TA, UK and ³Registre du Cancer de l'Isère, Centre Hospitalier Universitaire de Grenoble, BP 217, Pavillon E, 38043 Grenoble, Cedex 9, France

Background: There is considerable interest in the possibility of provision of lung cancer screening services in many developed countries. There is, however, no consensus on the target population or optimal screening regimen.

Methods: In this paper, we demonstrate the use of published results on lung cancer screening and natural history parameters to estimate the likely effects of annual and biennial screening programmes in different risk populations, in terms of deaths prevented and of human costs, including screening episodes, further investigation rates and overdiagnosis.

Results: Annual screening with the UK Lung Screening Study eligibility criteria was estimated to result in 956 lung cancer deaths prevented and 457 overdiagnosed cancers from 330 000 screening episodes. Biennial screening would result in 802 lung cancer deaths prevented and 383 overdiagnosed cancers for 180 000 screening episodes.

Interpretation/conclusion: The predictions suggest that the intervention effect could justify the human costs. The evidence base for low-dose CT screening for lung cancer pertains almost entirely to annual screening. The benefit of biennial screening is subject to additional uncertainty but the issue merits further empirical research.

Lung cancer is a major public health problem in the United Kingdom and worldwide. There were ~35 000 lung cancer deaths in the United Kingdom in 2010 (<http://www.cancerresearchuk.org/cancer-info/cancerstats/types/lung/>). Although there have been considerable improvements in prognosis of other cancers in recent decades, lung cancer is still characterised by poor survival, <10% at 5 years, partly because it tends to be diagnosed at a late stage (<http://seer.cancer.gov/statfacts/html/lungb.html>). Although the primary strategy for control of lung cancer is to reduce the prevalence of the main risk factor (i.e., smoking), it should be noted that many lung cancers occur in ex-smokers, and there may be a secondary role for early detection to diagnose the disease while it is successfully treatable (Raji *et al*, 2012). The National Lung Screening Trial (NLST) in the United States showed a 20%

reduction in lung cancer mortality with the offer of annual low-dose CT screening for lung cancer as compared with annual chest X-ray (Aberle *et al*, 2011). As a consequence, there is considerable interest in the possibility of a lung cancer screening service in the United Kingdom (Field and Duffy, 2008; Field *et al*, 2013). Lung cancer is unusual in that there is a single identified risk factor, cigarette smoking, that accounts for the majority of cases of the disease. This has two implications in terms of lung cancer screening: first, tobacco control activities should continue, with screening services linked to smoking cessation programmes; second, the target population for screening should be selected on the basis of estimation of risk of lung cancer, a strong component of which is smoking.

The NLST recruited subjects aged 55–74 with at least 30 pack-years of smoking history, including ex-smokers only if they

*Correspondence: Professor SW Duffy; E-mail: s.w.duffy@qmul.ac.uk

Received 30 August 2013; revised 13 January 2014; accepted 15 January 2014; published online 13 February 2014

© 2014 Cancer Research UK. All rights reserved 0007–0920/14

had given up during the past 15 years. The study group (26 722 individuals) was offered three annual screening episodes using low-dose CT over 2 years and the control group (26 732 individuals) three annual screens with chest X-ray. The outcome was based on lung cancer deaths within ~6 years from randomisation, using deaths from all lung cancers diagnosed after randomisation, that is, including cancers diagnosed in both arms after the screening stopped. There were 356 such deaths in the study arm and 443 in the control, a 20% relative reduction in the study arm, amounting to 87 deaths avoided in absolute terms.

The relative and absolute benefits would not necessarily be reproduced in a UK programme. The chest X-ray intervention in the control group may have prevented some deaths, and hence the effect of low-dose CT screening compared with usual care in the United Kingdom might be larger. A UK programme might target a different risk group that would affect the absolute baseline risk and potentially the absolute benefit. The compliance rates might be different (in the United Kingdom, there is ~75% compliance with breast screening and 60% with bowel screening). A future UK lung cancer screening programme might adopt a longer interscreening period than 1 year, which would imply a lower benefit, but would be likely to offer a longer period of screening, which might be expected to confer a higher absolute benefit.

The choice of which population to screen for lung cancer is an important one. Clearly, the target population should be at a notably higher risk than the general population. This is not simply a matter of cost effectiveness. In addition to the benefits, there are risks associated with screening, notably the procedures and anxiety associated with the investigation of suspicious imaging findings in those who do not have lung cancer (Seigneurin *et al*, 2013). Thus, the target population should be at a sufficiently high risk to yield a favourable benefit-harm balance of the intervention.

In this paper, we draw on the results of the NLST and of a number of other studies to estimate the likely effects, favourable and unfavourable, of a low-dose CT screening programme in the United Kingdom (Chien and Chen, 2008; Infante *et al*, 2009; Aberle *et al*, 2011; Pastorino *et al*, 2012; Saghir *et al*, 2012; Seigneurin *et al*, 2013; McRonald *et al*, 2014).

MATERIALS AND METHODS

Target populations, likely compliance and eligibility. We consider two possible risk groups: that recruited by NLST and that selected by the UK Lung Screening (UKLS) pilot trial. The former has empirically observed rates of lung cancer as reported in the trial, and the latter was chosen to have at least 5% risk of a diagnosis of lung cancer in the next 5 years, based on a validated risk model for lung cancer incorporating age, sex, smoking, family history, asbestos exposure and personal medical history (McRonald *et al*, 2014).

For compliance, the situation is slightly more complicated than in the usual screening programme, as the subjects must first be willing to supply the information necessary for risk assessment and, secondly, be willing to undergo screening should their personal risk meet the eligibility criteria. For simplicity, we combine these into a single compliance rate, of those willing to comply at both stages. We consider two possible scenarios: firstly, that this combined rate is 30%, as approximately observed in the UKLS pilot trial (Field *et al*, 2013) and, secondly, that following the publicity and heightened awareness that would accompany a national programme, a compliance rate of 60% would be observed.

However, it is unreasonable to expect that both populations would have the same proportions of eligible subjects. We therefore assume that the first of these, the 30% compliant, has a subgroup of 20% fulfilling the NLST criteria, that is, 6% of the overall

population, and 10% fulfilling the UKLS criteria (3% of the overall population, as approximately observed in UKLS). We assume that the 60% compliant have 16% eligible under the NLST criteria and 8% eligible under the UKLS (10 and 5% of the overall population). We assume that those who attend the prevalence screen also attend all scheduled incidence screens. There is a high correlation between successive screen attendance in other cancer screening programmes, but clearly the assumption of perfect correlation is at best an approximation.

The UKLS age range was 50–75 years and for NLST it was 55–74 years. Of the 30% compliant in UKLS, 10% satisfied the inclusion criteria. This may not necessarily generalise to the remaining 70%. It is estimated that ~15% of the US population in the age group of 55–74 years satisfy the NLST criteria (Ma *et al*, 2013).

Screen detection and interval cancer rates. The detection rate of cancers at screening is dependent on whether this is a first or subsequent screen, the incidence rate of disease in the population screened, the mean sojourn time (the average time spent in the presymptomatic, screen-detectable phase) and the sensitivity of the screening test (Weedon-Faekjer *et al*, 2010). If I denotes incidence, M the mean sojourn time and S the screening sensitivity, it has been shown (Launoy *et al*, 1998) that the expected prevalence of cancers detected at a first screen will be

$$P = IMS$$

For cancers detected at subsequent screens, the rate will additionally depend on the length of the interscreening interval. For a short interval, there will be relatively few interval cancers (cancers arising symptomatically between screens) and relatively few subsequent screen-detected cancers, as there is little time since the previous screen for new presymptomatic cancers to arise. For longer intervals, both interval cancer rates and subsequent screen detection rates will be higher. Calculation of these can be complicated, but Launoy *et al* (1998) showed that in steady state, in a population attending for screening, the expected proportion of cancers detected at subsequent screens (as opposed to arising in the interval between screens) can be estimated as

$$Q = \frac{S(1 - e^{-\lambda r})}{\lambda r(1 - (1 - S)e^{-\lambda r})}$$

where r is the interval between screens, $\lambda = 1/M$ is the transition rate from presymptomatic to symptomatic disease and S is the test sensitivity. This assumes an exponential distribution of the duration of the presymptomatic screen-detectable period. The expected rate of detection of cancers at a subsequent screen will be

$$P_1 = \frac{IrS(1 - e^{-\lambda r})}{\lambda r(1 - (1 - S)e^{-\lambda r})}$$

The expected absolute rate of cancers arising in the interval between two screens will be

$$I_1 = Ir \left(1 - \frac{S(1 - e^{-\lambda r})}{\lambda r(1 - (1 - S)e^{-\lambda r})} \right)$$

We take our estimates of S and M (and therefore λ) from Chien and Chen (2008), who estimated these quantities from a synthesis of published results from studies of lung cancer screening. We take two estimates of I , the underlying incidence in the population. For recruitment by the NLST criteria, we use the empirical annual incidence from the NLST (Aberle *et al*, 2011). For recruitment by the UKLS criteria, we use the estimated average incidence from the LLP risk model in 2848 positive responders in UKLS who met or exceeded the 5% risk criterion.

From the above, we estimate the numbers of cancers detected at screening and the numbers arising in the intervals between screens for different screening frequencies.

Further investigation rates. A proportion of those screened will have a suspicious finding, and will be recalled for further investigation. Of these, a further proportion will actually have lung cancer. The numbers of screenees recalled for further investigation will depend on the frequency of screening. For example, for some imaging findings, a repeat scan at 1 year is indicated. In an annual programme, this would not mean any additional diagnostic activity, whereas in a 2-yearly screening programme, it would mean at least one extra scan.

For annual screening, we estimate the numbers recalled for further investigation from NLST. Although in NLST positive findings were observed in 27%, 28% and 17% at first, second and third screens, respectively, not all of those required further diagnostic investigation, other than a repeat screen at 1 year. The percentage positive and requiring further investigation at first screen was 24%. The corresponding average percentage over the second and third screens was 13%. We therefore assume for annual screening a further investigation rate of 24% at first screen and 13% at subsequent. For 2-yearly screening, we assume 24% at all screens.

Effect on lung cancer mortality. We posit two possible absolute mortality effects. The first is simply a translation of the 20% reduction observed in NLST to an absolute effect. This was observed in the group offered CT screening in NLST, in association with three annual screens and high compliance rates. Although three other trials have published the effect on lung cancer mortality (Infante *et al*, 2009; Pastorino *et al*, 2012; Saghir *et al*, 2012) these trials were very small and a meta-analysis including all four trials gives a 19% reduction, very close to that of NLST (Field *et al*, 2013). We therefore use the NLST figure. The NLST figure may be conservative in any case, as the mortality includes deaths from tumours diagnosed after the first 3 years when both study and control arms were receiving usual care. In addition, the control arm in NLST was offered annual chest X-ray that may confer some mortality benefit, although a small one (Doria-Rose and Szabo, 2010; Oken *et al*, 2011). It is therefore both cautious and reasonable to assume a benefit of the order observed in NLST in association with annual CT screening. For brevity, we refer to this as the intent-to-treat (ITT) estimate as it is based on the randomised comparison in NLST. For 2-yearly screening, we assume the 20% to be attenuated to a 16% reduction, as the MST of 2.06 years and sensitivity of 97% would imply that the proportion of screen-detected cancers would be reduced by 20% with 2-yearly screening. It is not guaranteed that the mortality benefit would be attenuated by the same proportion as the number of screen-detected cancers, but the assumption would appear to be reasonable, firstly because the screening is only expected to change the prognosis of the screen-detected cancers and, secondly, because the phenomenon appears to hold in screening for other cancers, as evidenced by the differences by age in the proportions screen detected and the corresponding differences in mortality reductions in a breast screening trial (Tabar *et al*, 1992) and the differences between annual and biennial screening in a colorectal screening trial (Mandel *et al*, 1993).

The second estimate is based directly on the number of screen-detected cancers. The 20% reduction in lung cancer mortality in NLST corresponded to an absolute benefit of 13.4 deaths prevented per 100 screen-detected cancers (as there were 87 deaths prevented and 649 screen-detected cancers in the CT arm of NLST). The median follow-up time in NLST was 6.5 years, corresponding to an approximate 5.5 year follow-up of cancers from the time of diagnosis, and the fatality rate observed in the control group was 47%. With longer term follow-up of the order of 10 years, one

might reasonably expect at least 85% fatality as observed for 10-year fatality in SEER (<http://seer.cancer.gov/statfacts/html/lungb.html>). Our first estimate of potential lives saved for annual screening is derived by applying the 20% reduction to an expected combined fatality rate of 85% in the absence of screening from all cancers diagnosed in the programme. As noted above, with 47% case fatality in the control arm of NLST, there were 443 lung cancer deaths in the control arm and 356 in the CT arm. As 85% is 1.8 times 47%, we estimate that in the long term, there would be $1.8 \times 443 = 797$ deaths in the control arm of NLST and $1.8 \times 356 = 641$ deaths in the study arm, a difference of 156 deaths. Therefore, our second estimate of potential lives saved is 156 out of $649 = 24$ deaths prevented per 100 screen-detected cancers. We refer to this estimate as the per-protocol (PP) estimate.

Overdiagnosis. Overdiagnosis is usually defined as the diagnosis as a result of screening of cancer (generally histologically confirmed cancer) that would not have arisen in the host's lifetime if screening had not taken place. It is notoriously difficult to estimate, but an approximation can be obtained by considering that any excess incidence observed in the study group will be the sum of *overdiagnosed* cancers and *early-diagnosed* cancers, those whose diagnosis has been brought forward by the lead time conferred by screening. We therefore estimate the proportion of cancers with lead time that would extend beyond the period of observation. This in turn gives an estimate of the excess expected from lead time alone. This can be subtracted from the excess incidence observed in the study group to give an estimate of overdiagnosis. The proportion of screen-detected cancers that would be expected to have occurred after the period of observation in the absence of screening is in the notation

$$Q = e^{-5.5\lambda}$$

as the median time from diagnosis of screen-detected cancers to the end of follow-up was 5.5 years. This is calculated by substituting the estimate of λ from Chien and Chen (2008) and converted to an absolute expected excess in the study group. This is subtracted from the observed excess to give an estimate of overdiagnosis.

Uncertainty. Clearly, there are numerous sources of uncertainty in this estimation process. However, the uncertainty in the most important positive outcome, prevented deaths, is driven by the uncertainty in the mortality reduction of 20% observed in the NLST. We therefore estimated a range of uncertainty around our estimates of prevented deaths by applying the end points of the 95% confidence interval reported on this 20% (6.8–26.7%), and translating the end points of this interval to the estimated absolute numbers of deaths prevented.

Arguably, the most important negative end point is overdiagnosis, the estimate of which crucially depends on the estimated mean sojourn time of 2.06 years from the meta-analysis of Chien and Chen (2008). Accordingly, we estimated a range of uncertainties on the overdiagnosis based on the end points of the 95% confidence interval on this estimate (0.42–3.83 years).

RESULTS

Table 1 summarises screening and diagnostic results from the CT arm of NLST. There were 26 722 subjects randomised to the CT arm and 26 732 to the chest X-ray arm. In the average 6.5 years of observation, there were 1060 cancers diagnosed in the CT arm and 941 in the X-ray, a 13% excess (119 cancers) in the CT arm. On average, attendance at CT screening was 94%. The prevalence of lung cancers was 1.02% at first screen and 0.78% at the second and third screens combined.

Table 2 shows the estimated parameters from the procedures described in the Materials and Methods section. The risk criteria, eligibility, compliance and recall rates for further investigation are as explained above. The estimated mean sojourn time from the meta-analysis of Chien and Chen (2008) is 2.06 years, corresponding to $\lambda = 1/2.06 = 0.49$ under the assumption of exponential sojourn time. The same meta-analysis estimates sensitivity of low-dose CT screening as 97%. The annual incidences are as observed in NLST, 0.6%, and as predicted from the LLP risk score in UKLS, 1.4% (Aberle *et al*, 2011; McRonald *et al*, 2014). The case fatality in the absence of screening is estimated from the US SEER (<http://seer.cancer.gov/statfacts/html/lungb.html>). The absolute mortality reductions are estimated from the NLST results, both

directly and extrapolated to long term follow-up as described in the Materials and Methods section.

For overdiagnosis, we estimate that the proportion of screen-detected cancers in NLST brought forward from beyond the period of observation is

$$Q = e^{-5.5 \times 0.49} = 0.07$$

There were $270 + 168 + 211 = 649$ screen-detected cancers (Table 1). Thus, we would expect an excess of $0.07 \times 649 = 45$ cancers in the CT arm from lead time. In fact, the excess observed was 119, giving an estimated overdiagnosis of 74 cancers. Thus, we assume overdiagnosis of 11% (74 out of 649) of screen-detected cancers. The 95% confidence interval on the mean sojourn time of 0.42–3.83 years inverts to give a 95% interval on λ of 0.26–2.38, giving a range for Q of almost zero to 0.26. This in turn gives a range of overdiagnosis values from 0 to 18% of screen-detected cancers.

Table 3 shows the predicted outcomes based on the parameter estimates in Table 2. Predictions are shown for the eight combinations of the following: annual and biennial screening; UKLS and NLST populations (incidences of 1.4% and 0.6% per year respectively); and 30% and 60% compliance. The results are shown for an initial population of 1 million, which is successively reduced by excluding those ineligible on the basis of risk, and by noncompliance. Thus, the actual numbers screened are $\leq 10\%$ of the initial population. The interval cancers pertain to those actually screened. The deaths prevented are from those that would have occurred in an unscreened group within 10 years of diagnosis.

Table 4 summarises the results in terms of undesirable outcomes per death prevented. As expected, the outcomes for deaths prevented show a greater absolute benefit when using a higher-risk population. In annual screening of a population at UKLS risk, 1.4% per year, the ITT estimate gives 385 screens (330 000/857)

Table 1. Screening and diagnostic outcomes in the US National Lung Screening Trial

Detection status	Variable	Number
Screen 1	Persons screened	26 309
	Further investigations	6369
	Cancers	270
Screen 2	Persons screened	24 715
	Further investigations	3866
	Cancers	168
Screen 3	Persons screened	24 102
	Further investigations	2522
	Cancers	211
Between screens	Interval cancers	44
After screening ceased/nonattenders	Cancers diagnosed	367

Table 2. Parameters used in modelling outcomes of low-dose CT screening for lung cancer

Quantity	Source	Estimate	
		UKLS population	NLST population
Risk criterion	Aberle <i>et al</i> (2011), McRonald <i>et al</i> (2014)	$\geq 5\%$ in 5 years	≥ 30 pack-years
Compliance	McRonald <i>et al</i> (2014), hypothesised	30%, 60%	30%, 60%
Eligible (of total)	Aberle <i>et al</i> (2011), McRonald <i>et al</i> (2014)	3%, 5%	6%, 10%
MST (years)	Chien and Chen (2008)	2.06	2.06
CT Sensitivity	Chien and Chen (2008)	97%	97%
FI (screen 1)	Aberle <i>et al</i> (2011)	24%	24%
FI (screen 2+) annual	Aberle <i>et al</i> (2011)	13%	13%
FI (screen 2+) biennial	Aberle <i>et al</i> (2011)	24%	24%
Annual incidence	Aberle <i>et al</i> (2011), McRonald <i>et al</i> (2014)	1.4%	0.6%
Case fatality without screening	http://seer.cancer.gov/statfacts/html/lungb.html	85%	85%
Mortality reduction (annual)	Aberle <i>et al</i> (2011), extrapolated	20%, 27 out of 100 SDC	20%, 27 out of 100 SDC
Mortality reduction (biennial)	Aberle <i>et al</i> (2011), extrapolated	16%, 27 out of 100 SDC	16%, 27 out of 100 SDC
Overdiagnosis	Aberle <i>et al</i> (2011), Chien and Chen (2008)	0.11 per SDC	0.11 per SDC

Abbreviations: CT = computed tomography; FI = further diagnostic investigations; MST = mean sojourn time; NLST = National Lung Screening Trial; SDC = screen-detected cancer; UKLS = UK Lung Screening.

Table 3. Estimated outcomes per million population in the target age range, for 10 years of annual or biennial low-dose CT screening, based on the parameters in Table 2, with uncertainty intervals on the deaths prevented and overdiagnosed cancers

Screening frequency	Outcome	UKLS incidence Compliance		NLST incidence Compliance	
		30%	60%	30%	60%
Annual	Eligible and participating	30 000	50 000	60 000	100 000
	Prevalence screens	30 000	50 000	60 000	100 000
	Prevalence screen cancers	839	1399	719	1199
	Incidence screens	300 000	500 000	600 000	1 000 000
	Incidence screen cancers	3318	5530	2844	4740
	Interval cancers	882	1470	756	1260
	Further investigations	46 200	77 000	92 400	154 000
	Deaths prevented (ITT)	857 (291–1144)	1428 (486–1906)	734 (250–980)	1224 (416–1634)
	Deaths prevented (PP)	956 (325–1276)	1594 (542–2128)	819 (278–1093)	1366 (464–1824)
	Overdiagnosed cancers	457 (0–748)	733 (0–1247)	392 (0–641)	653 (0–1069)
Biennial	Eligible and participating	30 000	50 000	60 000	100 000
	Prevalence screens	30 000	50 000	60 000	100 000
	Prevalence screen cancers	839	1399	719	1199
	Incidence screens	150 000	250 000	300 000	500 000
	Incidence screen cancers	2646	4410	2268	3780
	Interval cancers	1554	2590	1332	2220
	Further investigations	43 200	72 000	86 400	144 000
	Deaths prevented (ITT)	685 (233–914)	1142 (388–1525)	587 (200–784)	979 (333–1307)
	Deaths prevented (PP)	802 (273–1071)	1336 (454–1784)	687 (234–917)	1145 (389–1529)
	Overdiagnosed cancers	383 (0–627)	639 (0–1046)	329 (0–538)	548 (0–896)

Abbreviations: CT = computed tomography; ITT = intent to treat; NLST = National Lung Screening Trial; PP = per protocol; UKLS = UK Lung Screening.

Table 4. Estimated outcomes and uncertainty intervals expressed as screening activity and undesired outcomes per lung cancer death prevented, stratified by incidence of target population, screening frequency and benefit estimate (ITT or PP)

Screening frequency, benefit estimate	Outcome per death prevented	UKLS incidence	NLST incidence
Annual, ITT	Persons screened for 10 years	35 (26–103)	82 (61–241)
	Screening episodes	385 (288–1132)	899 (673–2644)
	Further investigations	54 (40–159)	126 (94–371)
	Overdiagnosed cases	0.5 (0.0–0.9)	0.5 (0.0–0.9)
Annual, PP	Persons screened for 10 years	32 (24–94)	73 (55–215)
	Screening episodes	345 (258–1015)	805 (603–2368)
	Further investigations	48 (36–141)	113 (85–332)
	Overdiagnosed cases	0.4 (0.0–0.8)	0.4 (0.0–0.8)
Biennial, ITT	Persons screened for 10 years	44 (33–129)	102 (76–300)
	Screening episodes	263 (197–774)	613 (459–1803)
	Further investigations	63 (47–185)	147 (110–432)
	Overdiagnosed cases	0.6 (0.0–1.0)	0.6 (0.0–1.0)
Biennial, PP	Persons screened for 10 years	38 (28–112)	87 (65–256)
	Screening episodes	224 (168–659)	525 (393–1544)
	Further investigations	54 (40–159)	126 (94–371)
	Overdiagnosed cases	0.4 (0.0–0.8)	0.4 (0.0–0.8)

Abbreviations: ITT = intent to treat; NLST = National Lung Screening Trial; PP = per protocol; UKLS = UK Lung Screening.

and 54 subjects undergoing further investigation (46 200/857) per lung cancer death prevented. The corresponding figures for the NLST risk population, 0.6% per year, are 899 screens and 126 subjects undergoing further investigation per death prevented. The corresponding PP estimates are less conservative but show the same improved absolute benefit. Regardless of interval or target risk population, we estimate ~1 overdiagnosed case per 2 deaths prevented using ITT estimates and 1 overdiagnosed case per 2.5 deaths prevented using PP estimates.

Annual screening is estimated to be more effective in absolute terms, although the difference between the two regimens in terms of deaths prevented is modest. Annual screening incurs a lower cost per death prevented in terms of persons screened and in terms of numbers of further investigations, but a higher cost in terms of number of screening episodes. Absolute overdiagnosis rates are estimated to be lower in biennial screening, but rates of overdiagnosis per death prevented are the same in each regimen.

DISCUSSION

The above demonstrates the use of published figures to estimate the likely benefits and human costs of low-dose CT screening for lung cancer. The methodology is relatively simple, and the results are readily comprehensible. The predictions are based mainly on the largest published randomised trial of the intervention. We have included ranges of uncertainty on the numbers of deaths prevented and the overdiagnosed cases, based on published 95% confidence intervals on the mortality reduction and the mean sojourn time. However, one might consider the sensitivity of the outcomes to changes in the parameter estimates. If we assume the sensitivity recently reported by NLST of 93.8% (Church *et al*, 2013), instead of 97% from the meta-analysis, the numbers of screen-detected cancers and the numbers of deaths prevented would be attenuated by ~3%. More conservatively, use of the combined mortality reduction from all trials that have so far published lung cancer mortality results (Field *et al*, 2013) would reduce the numbers of deaths prevented by 5%. This would reduce the absolute numbers of deaths prevented by between 30 and 80. In any case, the results based on NLST are likely to be conservative, as noted above, because of the inclusion of post-screening cancers and the X-ray intervention in the control group. The fatality rate of the control group cases in NLST was considerably higher than one would expect from the SEER data (Aberle *et al*, 2011; <http://seer.cancer.gov/statfacts/html/lungb.html>). However, the range of possibilities from the alternative benefit estimates and postulated compliance is consistent with observations in a major review (Seigneurin *et al*, 2013).

Our own results here based on either NLST or UKLS are also likely to be conservative. In common with other modelling exercises to predict the effect of cancer screening, we have assumed an underlying Markov process model. This is a powerful tool for estimation and projection, but the Markov assumptions may not hold for all cancers. In this case, they appear to have underestimated incidence screen cancer detection rates, at 5 per 1000 screening episodes, in comparison with observed rates, 8 per 1000, in NLST (Aberle *et al*, 2011). Predicted interval cancer rates are correspondingly higher than observed in NLST. Thus, if anything, our PP estimates of lives saved are based on conservative estimates of the numbers of screen-detected cancers and therefore will be underestimates of the true numbers of deaths prevented.

We have not quoted uncertainty ranges for underlying incidence and mortality rates, or for rates of recall for further investigations, as these are generally based on very large numbers, and confidence intervals would consequently be narrow. It should, however, be acknowledged that incidence and mortality will

depend on the population targeted and further investigation rates on the diagnostic algorithm adopted, and to some extent on the hardware and software used in the CT scanning.

The predictions suggest that the intervention effect could justify the human costs. A full cost-effectiveness analysis from NLST is awaited. As one might expect, the use of a higher-risk population is a more cost-effective option. That is, for either annual or biennial screening, the screening is more efficient using the UKLS criteria for eligibility than using the NLST criteria. Against this, there are the National Health Service's considerations of equity and simplicity of the eligibility criteria.

Interestingly, the outcomes suggest that biennial screening, while being less effective in absolute terms than annual screening, may be similarly cost effective. The ITT estimate for the UKLS risk group is of 263 screening episodes and 63 further investigation episodes per death prevented, as compared with the figures of 385 and 54 above. However, it should be noted that the evidence base for low-dose CT screening for lung cancer pertains almost entirely to annual screening. The benefit of biennial screening was estimated by extrapolation from observed annual screening results, and hence is subject to additional uncertainty. This points up the limitations of modelling exercises, whether using simple deterministic approaches like this or more mathematical stochastic models (Chien and Chen, 2008; Weedon-Faekjer *et al*, 2010): the credibility of the model results rests crucially on the evidence base. For annual screening, there is already an experimental evidence base, whereas for biennial, the parameters have to be imputed from limited data. However, the results are sufficiently suggestive as to indicate that further empirical research on longer intervals than 1 year would be worthwhile.

The estimates of overdiagnosis suggest ~1 overdiagnosed case for every 2 to 2.5 lives saved. These too are subject to considerable uncertainty and further follow-up of the screening trials is necessary to obtain more reliable estimates.

REFERENCES

- Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, Fagerstrom RM, Gareen IF, Gatsonis C, Marcus PM, Sicks JD for the National Lung Screening Trial Research Team (2011) Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med* **365**: 395–409.
- Chien CR, Chen THH (2008) Mean sojourn time and effectiveness of mortality reduction for lung cancer screening with computed tomography. *Int J Cancer* **122**: 2594–2599.
- Church TR, Black WC, Aberle DR, Berg CD, Clingan KL, Duan F, Fagerstrom RM, Gareen IF, Gierada DS, Jones GC, Mahon I, Marcus PM, Sicks JD, Jain A, Baum S (2013) Results of initial low-dose computed tomographic screening for lung cancer. *NEJM* **368**: 1980–1991.
- Doria-Rose VP, Szabo E (2010) Screening and prevention of lung cancer. In *Lung Cancer: A Multidisciplinary Approach to Diagnosis and Management*, Kernstine KH, Reckamp KL (eds) pp 53–72. Demos Medical Publishing: New York.
- Field JK, Duffy SW (2008) Lung cancer screening: the way forward. *Br J Cancer* **99**: 557–562.
- Field JK, Hansell DM, Duffy SW, Baldwin DR (2013) Computed tomography screening for lung cancer: countdown to implementation? *Lancet Oncol* **14**(13): e591–e600.
- Infante M, Cavuto S, Lutman FR, Brambilla G, Chiesa G, Ceresoli G, Passera E, Angeli E, Chiarenza M, Aranzulla G, Cariboni U, Errico V, Inzillo F, Bottoni E, Voulaz E, Alloisio M, Destro A, Roncalli M, Santoro A, Ravasi G. for the DANTE Study Group (2009) A randomized study of lung cancer screening with spiral computed tomography: three-year results from the DANTE trial. *Am J Respir Crit Care Med* **180**: 445–453.
- Launoy G, Duffy SW, Prevost TC, Bouvier V (1998) Depistage des cancers: sensibilité du test et de la procédure de depistage. *Rev Epidemiologie Sante Publ* **46**: 420–426.

- Ma J, Ward EM, Smith R, Jemal A (2013) Annual number of lung cancer deaths potentially avertable by screening in the United States. *Cancer* **119**: 1381–1385.
- Mandel JS, Bond JH, Church TR, Snover DC, Bradley GM, Schuman LM, Ederer F (1993) Reducing mortality from colorectal cancer by screening for fecal occult blood. Minnesota Colon Cancer Control Study. *NEJM* **328**: 1365–1371.
- McRonal FE, Yadegarfar G, Baldwin DR, Devaraj A, Brain KE, Eisen T, Holemans JA, Ledson M, Screation NJ, Rintoul RC, Hands CJ, Lifford K, Whynes DK, Kerr KM, Page R, Parmar M, Wald N, Weller D, Williamson PR, Myles J, Hansell DM, Duffy SW, Field JK (2014) The UK Lung Screen (UKLS): demographic profile of first 88,897 approaches provides recommendations for population screening. *Cancer Prev Res (Phila)*; e-pub ahead of print 17 January 2014.
- Oken MM, Hocking WG, Kvale PA, Gerald L, Andriole GL, Buys SS, Church TR, Crawford ED, Fouad MN, Isaacs C, Reding DJ, Weissfeld JL, Yokochi LA, O'Brien B, Ragard LR, Rathmell JM, Riley TL, Wright P, Caparaso N, Ping H, Izmirlian G, Pinsky PF, Prorok PC, Barnett S, Kramer BS, Anthony B, Miller AB, Gohagan JK, Berg CD for the PLCO Project Team (2011) Screening by chest radiograph and lung cancer mortality: the Prostate, Lung, Colorectal, and Ovarian (PLCO) randomized trial. *JAMA* **306**: 1865–1873.
- Pastorino UA, Rossi MEF, Rosato VEF, Marchiano AB, Sverzellati NG, Morosi CB, Fabbri AC, Galeone CEF, Negri EE, Sozzi GD, Pelosi Giuseppe C, La Vecchia CEF, Carlo EF (2012) Annual or biennial CT screening versus observation in heavy smokers: 5-year results of the MILD trial. *Eur J Cancer Prev* **21**: 308–315.
- Raji OY, Duffy SW, Agbaje OF, Baker SG, Christiani DC, Cassidy A, Field JK (2012) Predictive accuracy of the Liverpool Lung Project Risk Model for stratifying patients for computed tomography screening for lung cancer: a case-control and cohort validation study. *Ann Intern Med* **157**: 242–250.
- Saghir Z, Dirksen A, Ashraf H, Bach KS, Brodersen J, Clementsen PF, Døssing M, Hansen H, Kofoed KF, Larsen KR, Mortensen J, Rasmussen JF, Seersholm N, Skov BG, Thorsen H, Tønnesen P, Pedersen JH (2012) 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* **67**: 296–301.
- Seigneurin A, Field J, Gachet A, Duffy S (2013) A systematic review of the characteristics associated with recall rates, detection rates and positive predictive values of computed tomography screening for lung cancer. *Ann Oncol*; e-pub ahead of print 1 December 2013.
- Tabar L, Fagerberg G, Duffy SW, Day NE, Gad A, Grøntoft O (1992) Update of the Swedish two-county program of mammographic screening for breast cancer. *Radiol Clin North Am* **30**: 187–210.
- Weedon-Faekjer H, Tretli S, Aalen OO (2010) Estimating screening test sensitivity and tumour progression using tumour size and time since previous screening. *Stat Methods Med Res* **19**: 507–527.

This work is published under the standard license to publish agreement. After 12 months the work will become freely available and the license terms will switch to a Creative Commons Attribution-NonCommercial-Share Alike 3.0 Unported License.