Annual or biennial CT screening versus observation in heavy smokers: 5-year results of the MILD trial

Ugo Pastorino[a], Marta Rossi[e,f], Valentina Rosato[e,f], Alfonso Marchianò[b], Nicola Sverzellati[g], Carlo Morosi[b], Alessandra Fabbri[c], Carlotta Galeone[e,f], Eva Negri[e], Gabriella Sozzi[d], Giuseppe Pelosi[c] and Carlo La Vecchia[e,f]

The efficacy and cost-effectiveness of low-dose spiral computed tomography (LDCT) screening in heavy smokers is currently under evaluation worldwide. Our screening program started with a pilot study on 1035 volunteers in Milan in 2000 and was followed up in 2005 by a randomized trial comparing annual or biennial LDCT with observation, named Multicentric Italian Lung Detection. This included 4099 participants, 1723 randomized to the control group, 1186 to biennial LDCT screening, and 1190 to annual LDCT screening. Follow-up was stopped in November 2011, with 9901 person-years for the pilot study and 17 621 person-years for Multicentric Italian Lung Detection. Forty-nine lung cancers were detected by LDCT (20 in biennial and 29 in the annual arm), of which 17 were identified at baseline examination; 63% were of stage I and 84% were surgically resectable. Stage distribution and resection rates were similar in the two LDCT arms. The cumulative 5-year lung cancer incidence rate was 311/100 000 in the control group, 457 in the biennial, and 620 in the annual LDCT group ($P=0.036$); lung cancer mortality rates were 109, 109, and 216/100 000 ($P=0.21$), and total mortality rates were 310, 363, and 558/100 000, respectively ($P=0.13$). Total mortality in the pilot study was similar to that observed in the annual LDCT arm at 5 years. There was no evidence of a protective effect of annual or biennial LDCT screening. Furthermore, a meta-analysis of the four published randomized trials showed similar overall mortality in the LDCT arms compared with the control arm. *European Journal of Cancer Prevention* **21:308–315 © 2012 Wolters Kluwer Health | Lippincott Williams & Wilkins.**

European Journal of Cancer Prevention 2012, **21**:308–315

Keywords: low-dose computed tomography, lung, neoplasms, randomized controlled trial, screening, smoking

[a]Unit of Thoracic Surgery, [b]Unit of Radiology, [c]Department of Pathology, [d]Unit of Tumor Genomics, Department of Experimental Oncology, Foundation IRCCS National Cancer Institute of Milan, [e]Department of Epidemiology, Mario Negri Institute for Pharmacological Research, [f]Department of Occupational Health, University of Milan, Milan and [g]Department of Clinical Sciences, Section of Radiology, University of Parma, Parma, Italy

Correspondence to Ugo Pastorino, MD, Unit of Thoracic Surgery, Foundation IRCCS National Cancer Institute of Milan, Via Venezian 1, 20133 Milan, Italy
Tel: +39 022 390 2906; fax: +39 022 390 2907;
e-mail: ugo.pastorino@istitutotumori.mi.it

Received 26 January 2012 **Accepted** 27 January 2012

Introduction

Lung cancer incidence and mortality have constantly declined during the last three decades in male populations of Europe and the US, mainly as a consequence of effective smoking control policies (Jemal *et al.*, 2010; La Vecchia *et al.*, 2010). This reduction is by far the most important determinant of the reduction in total cancer mortality observed for all sites. In the same period, however, the cure rates for lung cancer have not significantly improved, and the 5-year survival rate of all detected lung cancers remains below 15% (Verdecchia *et al.*, 2007).

At the turn of the century, the diagnostic potential of low-dose spiral computed tomography (LDCT) of the chest appeared to be a fundamental innovation, capable of radically changing the prospects of early lung cancer detection. In the pilot studies by Cornell University in New York, lung cancer screening with annual LDCT in heavy smokers was associated with a proportion of resectable stage I lung cancer and long-term survival rates in excess of 80% [Henschke *et al.*, 1999; The International Early Lung Cancer Action Program Investigators (I-ELCAP), 2006]. These results led the Early Lung Cancer Action Project (ELCAP) research group to estimate that LDCT screening in heavy smokers could prevent 80% of lung cancer mortality (I-ELCAP, 2006). A number of studies have been initiated since, involving over 60 000 individuals in single-arm and almost 90 000 in randomized trials, although with significant differences in the study design (Pastorino, 2010). The first two published randomized trials provided results that appeared below the expectations: a small Italian trial showed the same mortality in the CT screening and observation arms (Infante *et al.*, 2009), whereas the large US National Lung Screening Trial (NLST) including 53 454 individuals showed a 20% reduction in lung cancer mortality and a 7% reduction in total mortality in the 3-year annual CT arm compared with the annual chest radiograph arm (National Lung Screening Trial Research Team *et al.*, 2011).

Our screening program started in 2000 with a pilot study that enrolled 1035 volunteers to receive annual LDCT on a long-term basis, with selective use of PET. The early results at 3 years were encouraging and proved the additional diagnostic value of PET in a screening program

0959-8278 © 2012 Wolters Kluwer Health | Lippincott Williams & Wilkins DOI: 10.1097/CEJ.0b013e328351e1b6

(Pastorino *et al.*, 2003). In 2005, we launched the Multicentric Italian Lung Detection (MILD) study, a randomized trial comparing LDCT by annual or biennial intervals with observation (Marchianò *et al.*, 2009). The MILD trial also included extensive testing of blood and tissue biomarkers with predictive and early detection purposes.

We report here the 5-year follow-up results of the MILD trial and a historical comparison with the initial pilot study, in which the annual LDCT program was maintained up to 10 years.

Participants and methods

The Multicentric Italian Lung Detection project

The MILD project is an ongoing randomized lung cancer screening trial of which the primary aim is to evaluate the impact on mortality of early lung cancer detection through LDCT at annual or biennial intervals versus no screening, as well as promoting smoking cessation among participants and assessing the value of blood and tissue biomarkers in combination with LDCT (Marchianò *et al.*, 2009). The MILD project was initially designed as a multicentric trial, with a planned sample size of 10 000 individuals, a screening period of 10 years, and a total follow-up of 100 000 person-years. Such a sample size would be adequate to detect a 30% reduction in lung cancer mortality in the LDCT arm.

However, the national program faced many difficulties as a result of lack of funding, limited support from local authorities, and cultural prejudice: only a few hospitals from the Lombardy region obtained permission to start the trial, and recruitment was limited. For these reasons, the present analysis includes only the individuals enrolled and screened at the Istituto Nazionale Tumori of Milan. The study was approved by our Institutional Review Board and Ethics Committee. The sponsors had no role in conducting and interpreting the study.

Recruitment strategy, enrollment, and randomization of participants

Volunteers were recruited from among respondents to advertisements and articles published in the lay press and in television broadcasts. All volunteers were assessed for their eligibility by means of a questionnaire administered by entry phone/fax/e-mail or web. Eligibility criteria included age 49 years and above, current or former smokers (having quit smoking within 10 years of recruitment) with at least 20 pack-years of smoking, and no history of cancer within the previous 5 years.

Eligible participants were asked to read a detailed information sheet and sign a consent form. The participants were then randomly assigned to two groups: the control group underwent a program of primary prevention (smoking cessation) with pulmonary function test evaluation and blood sample collection, and the early detection group underwent the same program with the addition of LDCT. Centralized stratified randomization was accomplished by the use of blocks of variable size. The list of randomization was stratified by reference center, age (up to 65 years or older), and duration of smoking (more or less than 40 years). The group randomized to receive LDCT was further randomized to receive LDCT every 12 months (annual) or every 24 months (biennial).

Participants were then contacted to set up an appointment at the Early Detection Clinic, where they underwent clinical examinations according to the randomization arm and were administered a questionnaire including a detailed information on smoking history, personal and family medical history (cancer, cardiovascular, and respiratory diseases including chronic bronchitis, emphysema, and asthma), and information about attempts to stop smoking.

Low-dose spiral computed tomography and diagnostic workup

Multidetector CT was performed using a 16-detector-row CT system (Somatom Sensation 16; Siemens Medical Solutions, Forchheim, Germany). All CT examinations of the whole lung were conducted during one deep inspiratory breath-hold without the use of a contrast medium. The CT system was regularly calibrated to allow reliable measurements and comparison between examinations. Standard LDCT parameters were as follows: 120 kV, 30 mAs, 0.75 mm collimation, gantry rotation time 0.5 s, pitch 1.5. The MDCT data were reconstructed for the detection of pulmonary nodules. One-millimeter-thick sections with a reconstruction increment of 1 mm and a sharp kernel (Siemens B50 kernel, Siemens Medical Solutions) were used for the study. Axial, coronal, and oblique reformations (with a section width of 1 mm and a reconstruction increment of 1 mm) were independently evaluated by two trained radiologists, of whom one took the software-automated volume measurements (LungCare; Siemens Healthcare, Forchheim, Germany). In the event of a disagreement, a third radiologist was consulted.

Information on a maximum of four pulmonary nodules – excluding completely calcified nodules – was recorded for each CT examination. We considered solid lesions with a volume of less than 60 mm^3 (diameter of 4.8 mm or greater) nonsuspicious and scheduled repeat LDCT at 1 or 2 years. Nodules with a volume of 60–250 mm^3 (about 5–8 mm in diameter, respectively) underwent repeat CT examination after 3 months. Participants with nodules greater than 250 mm^3 were referred for additional workup, including fluorine 18-fluorodeoxyglucose PET or lung biopsy. We adopted a computer-aided detection volumetric growth of 25% or higher after a 3-month interval as the threshold indicative of malignant growth (Pastorino, 2010). No further evaluation was required until the next follow-up for nodules showing no growth.

Participants and follow-up

We enrolled a total of 4099 participants from September 2005 to January 2011. Of them, 1723 were randomized to the control group and 2376 to the LDCT group: 1190 were assigned to screening with annual LDCT and 1186 with biennial LDCT. They were followed up until November 2011 through active telephone follow-up and record linkage with the Cancer Registry Office database of Lombardy, which traced the vital status of all participants blindly, without knowing the random allocation.

For deceased participants, we obtained the death certificate from the Istituto Nazionale di Statistica (ISTAT). Information was missing from active telephone follow-up and Lombardy Cancer Registry Office database for two participants only, and they were censored at the last available date, thus leaving 4097 individuals (99.9%) for long-term analysis.

Historical comparison with the first pilot study

We also analyzed data from an update of the first prospective study started in 2000 at the European Institute of Oncology (IEO, pilot cohort). This included 1035 individuals aged 50 years or older who had smoked at least 20 pack-years (Pastorino *et al.*, 2003). All participants underwent annual LDCT (single-detector-row), with or without PET, for 10 years, and were followed up until January 2011 through record linkage with the Lombardy Cancer Registry Office. For three of them follow-up was incomplete, thus leaving 1032 individuals (99.7%) for long-term analysis.

Statistical analysis

Statistical analysis was performed at an independent research center, the Istituto Mario Negri of Milan. The primary endpoints were lung cancer and total mortality, and the secondary endpoints were lung cancer incidence and the number of procedures for benign lung diseases. Cumulative incidence of lung cancer and cumulative mortality from lung cancer and all causes were obtained as 1 minus the Kaplan–Meier estimator (Parmar and Machin, 1995). Differences among the three arms were assessed by means of the log-rank test. Hazard ratios (HR) and the corresponding 95% confidence intervals (CIs) were estimated using Cox proportional hazard models. All analyses were based on intention to screen.

Results

Table 1 gives the distribution of the three arms according to selected covariates at baseline. Age and gender were comparable in the three arms: 66.3% were men and 85.4% were younger than 65 years. The proportion of current smokers was higher in the control arm (89.7%) than in the LDCT arms (68.6%).

The median duration of follow-up was 4.4 years, with a maximum duration of 6 years for each group. The rate of adherence to the screening protocol was 95.1% in the biennial and 96.1% in the annual LDCT group.

A total of 9477 LDCTs were performed: 3763 in the biennial and 5714 in the annual LDCT arm. Participants who underwent at least one CT scan were 1149 in the biennial and 1152 in the annual LDCT group, with a median number of CTs of three in the biennial and five in the annual LDCT group. At baseline LDCT, 158 participants in the biennial and 177 in the annual LDCT group were considered positive or suspicious for lung cancer and required further evaluation, corresponding to a recall rate of 14 and 15%, respectively.

Thirty-four PETs were performed in the biennial and 49 in the annual LDCT group, corresponding to 2% of all participants and 0.9% of all LDCTs. Two participants in each group underwent surgery for detected benign nodules, representing 9% of all surgical procedures (4/45).

A total of 49 lung cancers were detected by LDCT screening, 20 in the biennial and 29 in the annual LDCT arm; 17 of them were detected at baseline examination (6 and 11 cases, respectively).

Table 2 shows selected characteristics of CT-detected lung cancers. Stage distribution and resection rates were similar in the two arms. Sixty-three percent of cancers were detected in stage I (70.0% in the biennial and 62% in the annual LDCT group, $P = 0.53$). The proportion of advanced disease (stage III–IV) was 25% in the biennial and 31% in the annual LDCT group. Resectability was 84% overall (85 vs. 83%, respectively), and the vast majority of patients were treated with lobectomy. Adenocarcinoma was the most frequent histology (65% overall; 85% in biennial and 52% in annual LDCT); other histologies, particularly squamous cell carcinoma, were more frequent in the annual than in the biennial LDCT arm (15 vs. 48%, $P = 0.016$).

A total of 20 lung cancers were diagnosed in the control group, 25 in the biennial and 34 in the annual LDCT groups (Table 3). The cumulative lung cancer incidence rate was 310.9/100 000 in the control group, 457.0 in the biennial, and 620.2 in the annual LDCT group. The number of interval lung cancers, not detected by screening, was the same in the two LDCT arms (five each). Deaths from lung cancer were 7, 6, and 12, respectively. Lung cancer mortality rates were 108.5/100 000 in the control, 108.8 in the biennial, and 216.0 in the annual LDCT groups. There were a total of 20 deaths in the control group as compared with 20 in the biennial and 31 in the annual LDCT groups. Total mortality rate was 310.1/100 000 in the control group, 362.5 in the biennial LDCT, and 557.9 in the annual LDCT.

Figure 1 shows the 5-year cumulative lung cancer incidence in the three arms. We found a significant difference between them ($P = 0.036$). In particular, there was an excess diagnosis in the LDCT groups in relation to the

Table 1 Selected baseline characteristics of 4099 Multicentric Italian Lung Detection participants by study arm

	Group [*N* (%)]		
	Control (*N*=1723)	Biennial CT (*N*=1186)	Annual CT (*N*=1190)
Age (years)			
<55	656 (38.1)	379 (32.0)	394 (33.1)
55–59	478 (27.7)	363 (30.6)	338 (28.4)
60–64	359 (20.8)	261 (22.0)	274 (23.0)
65–69	174 (10.1)	143 (12.1)	134 (11.3)
≥ 70	56 (3.3)	40 (3.4)	50 (4.2)
Median	57	58	57
Sex			
Male	1090 (63.3)	813 (68.5)	814 (68.4)
Female	633 (36.7)	373 (31.5)	376 (31.6)
Smoking status (smokers)			
Former smokers	177 (10.3)	376 (31.7)	370 (31.1)
Current smokers	1546 (89.7)	810 (68.3)	820 (68.9)
Duration of smoking (years)			
<30	140 (8.1)	98 (8.3)	102 (8.6)
30–39	856 (49.7)	584 (49.2)	604 (50.8)
40–49	619 (35.9)	442 (37.3)	412 (34.6)
≥ 50	108 (6.3)	62 (5.2)	72 (6.1)
Cigarettes per day (*N*)			
<20	568 (33.0)	282 (23.8)	262 (22.0)
20–29	731 (42.4)	614 (51.8)	619 (52.0)
30–39	241 (14.0)	141 (11.9)	142 (11.9)
≥ 40	183 (10.6)	149 (12.6)	167 (14.0)
Pack-years of cigarettes			
Median	38	39	39
FEV_1 (% predicted)[a]			
<90	330 (19.2)	328 (27.7)	336 (28.2)
≥ 90	1031 (59.8)	821 (69.2)	814 (68.4)

CT, computed tomography.
[a]The sum does not add to the total because of missing values.

Table 2 Selected characteristics of 49 computed tomography screen-detected lung cancers in the Multicentric Italian Lung Detection study during 5-year follow-up by the low-dose spiral computed tomography arm

	Group [*N* (%)]			
Lung cancers	Biennial CT (*N*=20)	Annual CT (*N*=29)	Total (*N*=49) [*N* (%)]	*P*-values
Sex				
Male	18 (90.0)	22 (75.9)	40 (81.6)	
Female	2 (10.0)	7 (24.1)	9 (18.4)	0.21
Histotype				
Carcinoma NOS	1 (5.0)	2 (6.9)	3 (6.1)	
Squamous cell carcinoma	1 (5.0)	10 (34.5)	11 (22.4)	
Adenocarcinoma[a]	17 (85.0)	15 (51.7)	32 (65.3)	
Large cell carcinoma	1 (5.0)	2 (6.9)	3 (6.1)	0.081
Stage				
IA	11 (55.0)	17 (58.6)	28 (57.1)	
IB	3 (15.0)	1 (3.4)	4 (8.2)	
IIA	–	1 (3.4)	1 (2.0)	
IIB	1 (5.0)	1 (3.4)	2 (4.1)	
IIIA	1 (5.0)	4 (13.8)	5 (10.2)	
IIIB	1 (5.0)	–	1 (2.0)	
IV	3 (15.0)	5 (17.2)	8 (16.3)	0.53
Resection				
None	3 (15.0)	5 (17.2)	8 (16.3)	
Lobectomy	16 (80.0)	20 (69.0)	36 (73.5)	
Segmentectomy	1 (5.0)	4 (13.8)	5 (10.2)	0.57

CT, computed tomography; NOS, not otherwise specified.
[a]*P*-value for $\chi^2_1=0.016$ for adenocarcinoma vs. others.

Table 3 Lung cancer incidence and mortality, and all-cause mortality per 100 000 person-years in the Multicentric Italian Lung Detection study at 5-year follow-up, by study arm

	Group					
	Control		Biennial CT		Annual CT	
	N	Rate	*N*	Rate	*N*	Rate
Person-years (incidence)	6432.9		5470.9		5481.9	
Person-years (mortality)	6449.5		5516.8		5556.7	
Lung cancer incidence	20	310.9	25	457.0	34	620.2
Lung cancer deaths	7	108.5	6	108.8	12	216.0
Total deaths	20	310.1	20	362.5	31	557.9

CT, computed tomography.

control group ($P = 0.025$) but not in the annual versus biennial LDCT groups ($P = 0.24$).

Figure 2 considers lung cancer mortality. There was no significant difference among arms ($P = 0.21$), and the HR was 1.52 (95% CI 0.63–3.65) when the two LDCT arms were together compared with the control group. After adjustment for age and smoking, the HR was 1.64 (95% CI, 0.67–4.01).

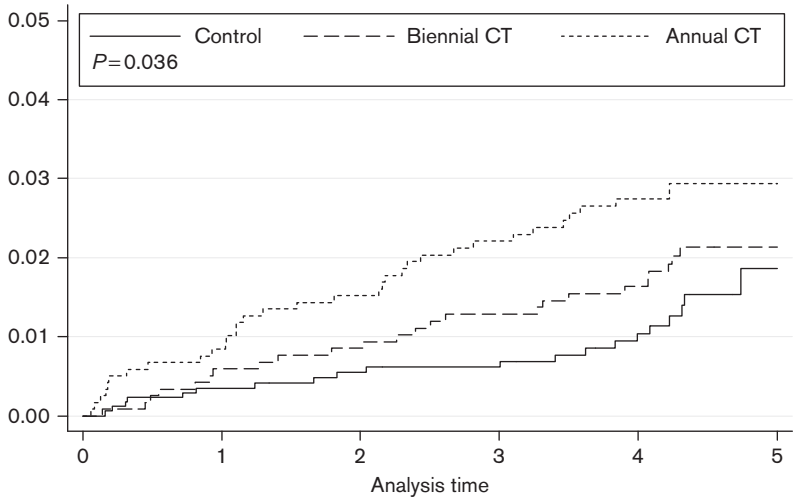
Figure 3 gives corresponding information for total mortality. There was no difference across groups ($P = 0.13$), and the HR was 1.39 (95% CI 0.83–2.34) when comparing the two LDCT arms together with the control group. After adjustment for age and smoking, the HR was 1.40 (95% CI, 0.82–2.38).

Figure 4 compares the cumulative total mortality between the annual LDCT arm of the MILD trial and the participants of the first pilot study conducted at the IEO, who underwent LDCT every year for 10 years and were followed up for 10 years for a total of 9901 person-years at risk. The two curves were virtually identical during the first 5 years of follow-up. The long-term follow-up of this pilot cohort showed a trend to increased mortality: from 387.6/100 000 at 1 year, to 1387.5/100 000 at 5 years, to 2590.0/100 000 at 10 years, consistent with aging of the cohort (from a median of 58 to 68 years).

Discussion

When the MILD trial was designed in 2004, there was little doubt among clinical oncologists about the efficacy of LDCT screening. A significant difference existed, however, in the estimated magnitude of the benefit, with speculations ranging from 20 to 80% reduction of lung cancer mortality to be expected as a consequence of systematic screening of heavy smokers. Flooded by information through newspapers and television, Italian smoking volunteers were eager to be tested by LDCT and reluctant to enter the control arm of any randomized trial. Thus, we had initially to propose a randomized comparison between two screening modalities: annual versus biennial LDCT. Only when this study had been approved and funded did we put forward the addition of an observational control arm. This explains the lower

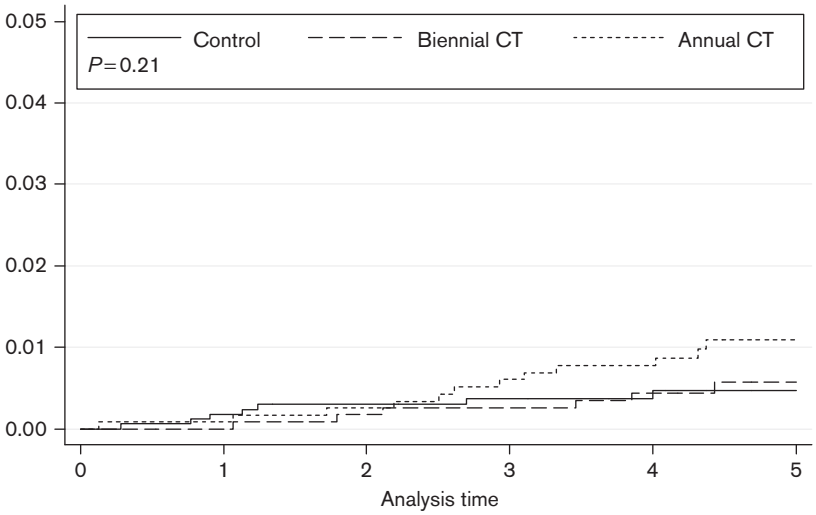
Fig. 1



Number at risk						
Control	1723	1678	1460	1336	1027	120
Biennial CT	1186	1172	1152	1131	1076	386
Annual CT	1190	1177	1147	1117	1065	462

Cumulative lung cancer incidence for the control group, biennial group, and annual low-dose computed tomography (CT) groups of the Multicentric Italian Lung Detection study during 5-year follow-up.

Fig. 2

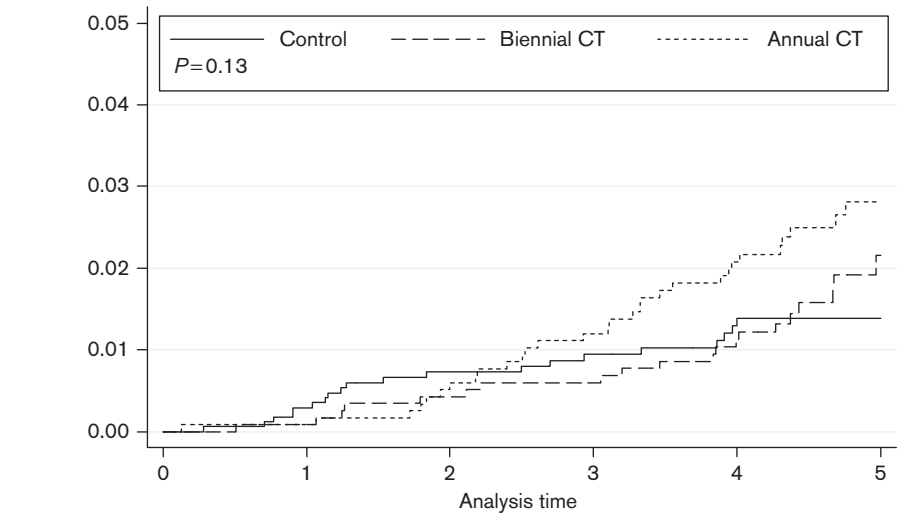


Number at risk						
Control	1723	1681	1464	1340	1032	121
Biennial CT	1186	1179	1160	1143	1089	394
Annual CT	1190	1185	1161	1135	1086	473

Cumulative lung cancer mortality for the control group, biennial group, and annual LD computed tomography (CT) groups of the Multicentric Italian Lung Detection study during 5-year follow-up.

number in the control group. The debate was so strong that it took a few months and four consecutive meetings of our institutional Ethics Committee to get the new design approved in 2005. All these elements together jeopardized the development and funding of the planned Italian multicentric trial.

Fig. 3

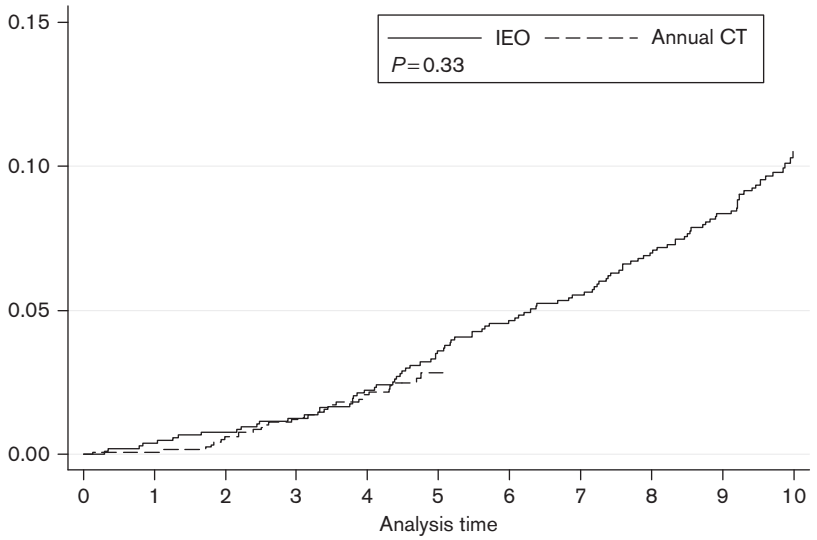


Number at risk

Control	1723	1681	1464	1340	1032	121
Biennial CT	1186	1179	1160	1143	1089	394
Annual CT	1190	1185	1161	1135	1086	473

Cumulative all-cause mortality for the control group, biennial group, and annual LD computed tomography (CT) groups of the Multicentric Italian Lung Detection study during 5-year follow-up.

Fig. 4



Number at risk

IEO	1032	1028	1024	1019	1009	995	984	975	952	938	434
Annual CT	1190	1185	1161	1135	1086	473					

Cumulative all-cause mortality for the European Institute of Oncology (IEO) cohort and the annual LDCT group of the Multicentric Italian Lung Detection study during 10-year follow-up.

Since then, the perception has substantially changed, and most clinicians are now convinced that the benefit of LDCT screening, if real, is small (Bach *et al.*, 2007; Bach, 2008). In this view, the MILD trial now appears underpowered and unable to detect differences in the order of 10% or less. Nevertheless, there are

peculiar aspects and useful information in this trial that cannot be provided by any of the other ongoing studies.

First, the MILD trial confirmed the value of selective use of PET, in addition to automated measurement of volume changes through the Lung-Care software, for the diagnostic workup of suspicious nodules (Pastorino *et al.*, 2009). In fact, one of the major side effects of LDCT screening, that is, the frequency of surgical procedures for benign disease, was only 9% in the MILD trial, compared with 27% in the concurrent Nelson trial, which applied a diagnostic algorithm based only on volumetric assessment of growth at LDCT (Van Klaveren *et al.*, 2009).

The second element consists of the relative efficacy of a different screening intensity. Our observation that lung cancer incidence was higher in the annual than in the biennial LDCT arm, without any shift to higher stages in the biennial arm (which could be anticipated in the case of late detection), provides additional evidence of overdiagnosis. The discrepancy in lung cancer incidence among the three arms was somewhat mitigated by our policy of conservative management of ground-glass opacities (GGOs). In fact, in the MILD trial, pure GGOs have never been biopsied or resected, despite increase in size, unless they developed a solid component within the lesion. This policy may also explain the lower frequency of adenocarcinomas (65%) and the extreme rarity of in-situ adenocarcinoma (former bronchioloalveolar carcinoma) in the MILD trial compared with other LDCT screening experiences (Vazquez *et al.*, 2008; Pastorino, 2010). It is noteworthy that no advanced lung cancer, interval cancer, or cancer death occurred in the group of participants with GGO. The lower incidence of squamous carcinomas in the biennial LDCT arm (1 vs. 10 cases, $P = 0.015$) also suggests that some of these lesions undergo spontaneous regression if not resected.

Even though the total number of deaths is small, and none of the observed differences reaches statistical significance, lung cancer and total mortality were, if anything, apparently higher in the annual LDCT arm compared with the control arm. Similarly, the 5-year results of the Danish trial (DLCST) have been presented recently, comparing five rounds of annual LDCT versus observation in a similar population (4104 individuals) (Saghir *et al.*, 2011). That trial did not find lower lung cancer mortality in the LDCT arm (15 vs. 11, $P = 0.428$) and showed a borderline significant excess of total mortality (61 vs. 42, $P = 0.059$). If the results of the MILD and DLCST trials are pooled together, the detrimental effect of LDCT screening on total mortality becomes statistically significant (112 vs. 62 deaths, $P = 0.005$). Moreover, the protective effect of LDCT on total mortality, shown by the NLST trial, disappears with a pooled analysis of the four published trials [Dante (Infante *et al.*, 2009), NLST (National Lung Screening Trial Research Team *et al.*, 2011), MILD, and DLCST (Saghir *et al.*, 2011); 2028 vs. 2083 deaths, relative risk 0.95, 95% CI 0.89–1.02)]. The effect on lung cancer mortality remains significant (relative risk 0.82, 95% CI 0.73–0.93) but the value of disease-specific mortality as the only endpoint appears questionable for two reasons: the assessment of the real cause of death can be very difficult in heavy smokers because of complex comorbidity; a shift in the cause of death from one disease to another is frequent in screened populations and hence potentially misleading. In any case, this would be a trivial goal in a population in which primary prevention (smoking cessation) would be effective against most causes of death.

A third piece of original information is related to the duration of screening, captured by the long-term analysis of our pilot trial started in 2000. To our knowledge, this is the first study in which LDCT was performed annually for 10 years and follow-up information was obtained for 99.7% of participants. The equivalence in 5-year mortality of the pilot study and the annual LDCT arm of the MILD trial is reassuring and excludes any potential selection bias as a cause of the higher mortality observed in the MILD population. Further, the slope of the whole curve is more important in the assessment of screening efficacy. In fact, it has been postulated by the ELCAP research team that a 3-year intervention plan, such as in the NLST trial, would determine an initial fall in mortality, followed by a new increase in mortality after suspension of screening (Henschke *et al.*, 2011). The constant increase in mortality, in the absence of any change in slope, that we observed in our prolonged screening cohort is against such a hypothesis and does not suggest any protective effect of long-term screening. In fact, mortality increased from 0.4% in the first year to 2.6% in the 10th year, despite annual LDCT investigation throughout this period, and the observed cumulative mortality rate of 10% at 10 years is comparable to the expected and for this cohort in the absence of screening.

The NLST trial has been successful in terms of recruitment rate, compliance, and quality of participating centers. Future analyses will clarify important elements such as side effects and costs of screening and impact of current smoking on the final outcome. All these data are crucial for the evaluation of screening in a preventable disease. There are, however, two limiting factors to be considered: (a) the trial does not include a proper control arm; instead, it compares two screening modalities such as 3-year annual LDCT versus 3-year chest radiograph; (b) early stopping of the trial may well overestimate the real effect of intervention (Bassler *et al.*, 2010).

The European trials have a proper observational control arm but are underpowered. The only way to circumvent this problem will be a pooled analysis that would reach the reasonable size of 33 000 individuals and might be feasible in less than 2 years from now.

Beyond the specific question of mortality reduction, a decade of clinical research on LDCT screening has markedly changed our knowledge of the natural history

and biology of lung cancer. In fact, collateral studies of the MILD project have provided new insight into the genetic determinants of tobacco addiction (Falvella *et al.*, 2010), chronic obstructive pulmonary disease (Calabrò *et al.*, 2010), and coronary calcification (Sverzellati *et al.*, 2012) as independent risk factors for lung cancer, frequency of interstitial lung disease (Sverzellati *et al.*, 2011) and bronchial diverticula (Sverzellati *et al.*, 2010), and the value of tissue and blood biomarkers (Roz *et al.*, 2009; Sozzi *et al.*, 2009; Cremona *et al.*, 2010). After 15 years of extensive research on circulating DNA (Sozzi *et al.*, 1999, 2003, 2005), we could demonstrate that microRNA signatures in plasma can not only detect lung cancer 2 years earlier than LDCT, but also predict the aggressiveness of disease and distinguish indolent from lethal cancers (Boeri *et al.*, 2011). Such a discovery will help clarify why the most virulent forms of lung cancer elude LDCT screening (Early warnings, 2009), and will open new perspectives in the early detection and management of lung cancer (Sozzi *et al.*, 2011).

Acknowledgements

The authors thank Claudio Jacomelli for data management, Tiziana Codazzi of the Lombardy Cancer Registry, and the MILD secretariat (Elena Bertocchi, Carolina Ninni, Annamaria Calanca).

This project was supported by a research grant from the Italian Association for Cancer Research (AIRC), the Italian Ministry of Health, the Lombardy Region, and the Cariplo Foundation.

Conflicts of interest

There are no conflicts of interest.

References

Bach PB (2008). Is our natural-history model of lung cancer wrong? *Lancet Oncol* **9**:693–697.

Bach PB, Jett JR, Pastorino U, Tockman MS, Swensen SJ, Begg CB (2007). Computed tomography screening and lung cancer outcomes. *JAMA* **297**:953–961.

Bassler D, Briel M, Montori VM, Lane M, Glasziou P, Zhou Q, *et al.* (2010). Stopping randomized trials early for benefit and estimation of treatment effects: systematic review and meta-regression analysis. *JAMA* **303**:1180–1187.

Boeri M, Verri C, Conte D, Roz L, Modena P, Facchinetti F, *et al.* (2011). MicroRNA signatures in tissues and plasma predict development and prognosis of computed tomography detected lung cancer. *Proc Natl Acad Sci USA* **108**:3713–3718.

Calabrò E, Randi G, La Vecchia C, Sverzellati N, Marchianò A, Villani M, *et al.* (2010). Lung function predicts lung cancer risk in smokers: a tool for targeting screening programmes. *Eur Respir J* **35**:146–151.

Cremona M, Calabrò E, Randi G, De Bortoli M, Mondellini P, Verri C, *et al.* (2010). Elevated levels of the acute-phase serum amyloid are associated with heightened lung cancer risk. *Cancer* **116**:1326–1335.

Early warnings (2009). Screening programmes for cancer detection are not always as effective at saving lives as might be hoped. Improving the situation will require a concerted effort on a broad front (editorial). *Nature* **458**:679.

Falvella FS, Galvan A, Colombo F, Frullanti E, Pastorino U, Dragani TA (2010). Promoter polymorphisms and transcript levels of nicotinic receptor CHRNA5. *J Natl Cancer Inst* **102**:1366–1370.

Henschke CI, McCauley DI, Yankelevitz DF, Naidich DP, McGuinness G, Miettinen OS, *et al.* (1999). Early Lung Cancer Action Project: overall design and findings from baseline screening. *Lancet* **354**:99–105.

Henschke CI, Boffetta P, Gorlova O, Yip R, DeLancey JO, Foy M (2011). Assessment of lung-cancer mortality reduction from CT screening. *Lung Cancer* **71**:328–332.

Infante M, Cavuto S, Lutman FR, Brambilla G, Chiesa G, Ceresoli G, *et al.*; 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 Resp Crit Care Med* **180**:445–453.

Jemal A, Siegel R, Xu J, Ward E (2010). Cancer statistics, 2010. *CA Cancer J Clin* **60**:277–300.

La Vecchia C, Bosetti C, Lucchini F, Bertuccio P, Negri E, Boyle P, *et al.* (2010). Cancer mortality in Europe, 2000–2004. *Ann Oncol* **21**:1323–1360.

Marchianò A, Calabrò E, Civelli E, Di Tolla G, Frigerio LF, Morosi C, *et al.* (2009). Pulmonary nodules: volume repeatability at multidetector CT lung cancer screening. *Radiology* **251**:919–925.

National Lung Screening Trial Research Team, Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, Fagerstrom RM, *et al.* (2011). Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med* **365**:395–409.

Parmar MKB, Machin D (1995). *Survival analysis: a practical approach*. Chichester: Wiley; p. 255.

Pastorino U (2010). Lung cancer screening. *Br J Cancer* **2**:1681–1686.

Pastorino U, Bellomi M, Landoni C, De Fiori E, Arnaldi P, Picchio M, *et al.* (2003). Early lung-cancer detection with spiral CT and positron emission tomography in heavy smokers: 2-year results. *Lancet* **362**:593–597.

Pastorino U, Landoni C, Marchianò A, Calabrò E, Sozzi G, Miceli R, *et al.* (2009). Fluorodeoxyglucose (FDG) uptake measured by positron emission tomography (PET) and standardised uptake value (SUV) predicts long-term survival of CT screening-detected lung cancer in heavy smokers. *J Thorac Oncol* **11**:1352–1356.

Roz L, Verri C, Conte D, Miceli R, Mariani L, Calabrò E, *et al.* (2009). Plasma DNA levels in spiral CT-detected and clinically detected lung cancer patients: a validation analysis. *Lung Cancer* **66**:270–271.

Saghir Z, Dirksen A, Ashraf HG, Tønnesen P, Bach KS, Hansen H, *et al.* (2011). The Danish Lung Cancer Screening Trial (DLCST): status after five years of CT screening. Proceedings of the World Conference on Lung Cancer, Amsterdam, July 2011. Abstract O17.03.

Sozzi G, Musso K, Ratcliffe C, Goldstraw P, Pierotti MA, Pastorino U (1999). Detection of microsatellite alterations in plasma DNA of non-small cell lung cancer patients: a prospect for early diagnosis. *Clin Cancer Res* **5**:2689–2692.

Sozzi G, Conte D, Leon M, Ciricione R, Roz L, Ratcliffe C, *et al.* (2003). Quantification of free circulating DNA as a diagnostic marker in lung cancer. *J Clin Oncol* **21**:3902–3908.

Sozzi G, Roz L, Conte D, Mariani L, Andriani F, Verderio P, *et al.* (2005). Effects of prolonged storage of whole plasma or isolated plasma DNA on the results of circulating DNA quantification assays. *J Natl Cancer Inst* **97**:1848–1850.

Sozzi G, Roz L, Conte D, Mariani L, Andriani F, Lo Vullo S, *et al.* (2009). Plasma DNA quantification in lung cancer computed tomography screening: five-year results of a prospective study. *Am J Respir Crit Care Med* **179**:69–74.

Sozzi G, Pastorino U, Croce CM (2011). MicroRNAs and lung cancer: from markers to targets. *Cell Cycle* **10**:2045–2046.

Sverzellati N, Ingegnoli A, Calabrò E, Randi G, La Vecchia C, Marchianò A, *et al.* (2010). Bronchial diverticula in smokers on thin-section CT. *Eur Radiol* **20**:88–94.

Sverzellati N, Guerci L, Randi G, Calabrò E, La Vecchia C, Marchianò A, *et al.* (2011). Interstitial lung diseases in a lung cancer screening trial. *Eur Respir J* **38**:392–400.

Sverzellati N, Cademartiri F, Bravi F, Martini C, Gira FA, Maffei E, *et al.* (2012). Relationship and prognostic value of modified coronary artery calcium score, FEV1, and emphysema in lung cancer screening population: The MILD Trial. *Radiology* **262**:460–467.

The International Early Lung Cancer Action Program Investigators (I-ELCAP) (2006). Survival of patients with stage I lung cancer detected on CT screening. *N Engl J Med* **355**:1763–1771.

Van Klaveren RJ, Oudkerk M, Prokop M, Scholten ET, Nackaerts K, Vernhout R, *et al.* (2009). Management of lung nodules detected by volume CT scanning. *N Engl J Med* **361**:2221–2229.

Vazquez M, Carter D, Brambilla E, Gazdar A, Noguchi M, Travis WD, *et al.* (2008). The International Early Lung Cancer Action Program Investigators. Solitary and multiple resected adenocarcinomas after CT screening for lung cancer: histopathologic features and their prognostic implications. *Lung Cancer* **64**:148–154.

Verdecchia A, Franceschi S, Brenner H, Gatta G, Micheli A, Mangone L, *et al.* (2007). Recent cancer survival in Europe: a 2000–02 period analysis of EUROCARE 4 data. *Lancet Oncol* **8**:784–796.