Self-Selection Effects in Smokers Attending Lung Cancer Screening

A 9.5-Year Population-Based Cohort Study in Varese, Italy

Lorenzo Dominioni, MD, Nicola Rotolo, MD,* Albino Poli, BS,† Massimo Paolucci, MD,‡ Fausto Sessa, MD,§ Vincenzo D'Ambrosio, MD,‖ Antonio Paddeu, MD,¶ William Mantovani, MD,† and Andrea Imperatori, MD**

Background: We hypothesize that mortality risk profile of participants and nonparticipants in nonrandomized lung cancer (LC) screening of smokers may be different.

Methods: In 1997, a population-based cohort of 5815 smokers of Varese Province was invited to nonrandomized LC screening by annual chest x-ray examination for 4 years. LC risk factors and screening participation rate were recorded. Except for screening, the whole cohort received usual care. After 9.5-year observation, we compared mortality of participants versus nonparticipants by assessing age-standardized all-cause mortality rate ratio (MRR) and disease group-specific MRR with 95% confidence intervals (95% CI).

Results: Self-selected screening participants were 21% of cohort. Participants were younger ($p < 0.001$), were more frequently current smokers ($p = 0.019$), had more pack-years of smoking ($p < 0.0001$), and had higher rate of LC family history ($p < 0.0001$) and of occupational LC risk ($p < 0.0001$) relative to nonparticipants. In logistic regression analysis familial LC, occupational risk and pack-years smoked were significant predictors of participation in screening and of developing LC. Participants displayed a healthy effect, as shown by all-cause MRR = 0.67 (95% CI, 0.53–0.84), all cancers except LC MRR = 0.61 (95% CI, 0.41–0.91), cardiovascular diseases MRR = 0.38 (95% CI, 0.22–0.63), and noncancer disease other than cardiovascular or respiratory MRR = 0.57 (95% CI, 0.34–0.92). The LC mortality (MRR = 1.40; 95% CI, 1.03–1.91) was higher in participants relative to nonparticipants ($p = 0.031$).

Conclusion: The selection effect in LC screening participants was dual: healthy effect and higher LC mortality. In assessing the overall effectiveness of LC screening on a population level, a higher LC mortality risk in participants should be considered.

Key Words: Lung cancer screening, self-selection, population-based study.

(*J Thorac Oncol.* 2010;5: 428–435)

Selection of volunteer participants, characterized by the "healthy effect," has been shown in screening, and it has raised concern about the generalizability to the population of interest of the results of randomized cancer screenings.[1–5] In lung cancer (LC) screening of volunteers, the possible effects of self-selection on mortality risk have been postulated[6,7] but have never been studied. Differences of mortality profile between screening participants and nonparticipants may be relevant for the interpretation of results of nonrandomized LC screening on a population level.

The aim of this study was to estimate the age-standardized all-cause mortality rate ratio (MRR) and the disease group-specific MRR for participants relative to nonparticipants in the PREcoce DIagnosi CAncro-early diagnosis of cancer (PREDICA) project, a nonrandomized LC screening study, LC screening or usual care, offered to a cohort of smokers in Varese, Italy.

METHODS

The PREDICA Project

A pilot study of LC screening, named PREDICA project, is ongoing since 1997 in a population-based cohort of previously unscreened smokers of Varese, Italy.[8] The screening protocol consisted of baseline two-view chest x-rays (CXRs) and annual repeat CXR for 4 years, with 5-year follow-up after the end of the enrollment period (January 1, 2002, to December 31, 2006). It was offered free of charge to heavy or long-term smokers (hereafter referred to as "PREDICA cohort") possessing the following eligibility criteria: asymptomatic cigarette smoker of more than 10 pack-years (current smoker or ex-smoker for <10 years), resident in Varese Province, both genders, ages 45 to 75 years, fit for

*Center for Thoracic Surgery, Department of Surgical Sciences, University of Insubria, Varese, Italy; †Department of Medicine and Public Health, University of Verona, Verona, Italy; ‡Department of Radiology, Ospedale S. Antonio Abate, Gallarate, Italy; §Department of Human Morphology, University of Insubria, Varese, Italy; ‖Thoracic Medicine Unit, Department of Medicine, Ospedale S. Antonio Abate, Gallarate, Italy; and ¶Respiratory Care Unit, Department of Medicine, Ospedale S. Anna, Como, Italy.

Disclosure: The authors declare no conflicts of interest.

Address for correspondence: Lorenzo Dominioni, MD, Center for Thoracic Surgery, University of Insubria, Ospedale di Circolo, Viale Borri 57, 21100 Varese, Italy. E-mail: lorenzo.dominioni@uninsubria.it

ISSN: 1556-0864/10/0504-0428

possible thoracotomy, and without diagnosed or suspected LC. No other exclusion criteria were used.

The PREDICA Cohort

The PREDICA cohort was recruited among all residents registered in the medical practices of 50 physicians public general practitioners (GPs) collaborating on this project (Appendix A). These practices, scattered over rural and urban areas in 44 towns of the Varese Province (hereafter referred to as "PREDICA towns," listed in Appendix B) included about 60,000 adults. In the whole Varese Province, there are 805,000 inhabitants served by 597 public GPs. The latter are paid with National Health System public funds to provide free medical care to all the resident population. The proportion of the Varese Province population between the age 45 and 75 years served by the 50 collaborating GPs is 8.5%. In early 1997, the 50 GPs compiled a preliminary recruitment list including 5925 smokers, encrypted to comply with privacy regulations, which pooled all resident smokers aged 45 to 75 years potential participants in screening. After excluding 110 noneligible candidates (unfit for surgery or with diagnosed/suspected LC), the PREDICA cohort comprised 5815 smokers, all whites. Of the latter, we recorded age, gender, smoker status (current/former smoker), and education level (illiterate, elementary school, intermediate school, high school, and university). Moreover, three LC risk factors were recorded: number of pack-years of smoking, family history of LC in first-degree relatives,[9] and occupational exposure to potential carcinogens.[10]

Until the PREDICA cohort recruitment list was complete, targeted subjects were told nothing of the study. In May 1997, an invitation letter to participate in LC screening was mailed to the 5815 cohort smokers, informing about risks and possible benefits of screening (significantly increased resectability and survivorship of screening-detected LC).

CXR Screening and Follow-Up

Individuals who accepted to be screened signed informed consent and were enrolled by their GPs. For nonparticipants, the informed consent was waived. The usual standard of care provided by the National Health Service GPs was the same in the whole PREDICA cohort, with the only addition of annual CXR in screening participants. After undertaking baseline CXR examination, the participants with negative CXR reading were scheduled for subsequent annual CXR appointments. Screening CXRs were taken in two views (posteroanterior: 100–120 KV, 200 mA, 6–8 mAs; lateral: 120–130 KV, 200 mA, 7–9 mAs; mean total effective dose: 1.6 mGy) and were interpreted by one of three senior thoracic radiologists. Uncertain and cancer-suspicious CXR diagnoses were referred to our team of thoracic oncologists. All LCs diagnosed, by screening or outside screening, were treated by usual international criteria.[11] Candidates to active LC treatment had histologic/cytologic confirmation of diagnosis.

Enrolment and active screening period lasted 4.5 years, from July 1, 1997, until December 31, 2001; the 5815 cohort subjects were allocated either as participant or nonparticipant progressively throughout that period. Follow-up was extended for 5 more years, until December 31, 2006, and the total observation period of the cohort was 9.5 years. For all cohort subjects, the person-years of observation were computed as the time between start of study and date of study cutoff or date of death, whichever came first, regardless of date of allocation as participant or nonparticipant. Given the study design, there were no dropouts: subjects who undertook at least the baseline CXR screening by December 31, 2001, were allocated as participants; all other subjects of the cohort were allocated as nonparticipants.

Incidence of LC

All new LCs diagnosed in the cohort were recorded, adopting the Varese Cancer Registry criteria,[12] by linkage with Varese Epidemiology Observatory,[13] Varese Province hospitals records and pathology records. For each new LC, we recorded date of diagnosis, whether diagnosis was clinicoradiologic or histologic, and histotype when available. All LCs diagnosed during the study (at baseline or annual screening, interval cases, and cases detected outside screening) were defined "incident" LCs.

Mortality

Deaths occurred in the cohort during the 9.5-year observation were searched by linkage with the Varese Mortality Registry[13]; date and cause of death were recorded. The cohort subjects who did not enroll in screening or who died before end of enrolment were arbitrarily classified as nonparticipants. Deaths attributable to complications of treatment of LC were filed as deaths from LC. The Varese Mortality Registry was accessed in 2008, completed with certificates of deaths occurred until December 31, 2006. Death certificates were reviewed by the mortality review committee (L.D., A.I., N.R., F.S., A.Po., and W.M.); only deaths definitely attributed to LC were reported as "LC deaths" in the results. Sensitivity of death certificates in reporting LC as cause of death was measured using the cause of death attributed by the mortality review committee as the gold standard; sensitivity was computed as the proportion of LC death certificates concordant with the mortality review committee's assessment of death attributed to LC.[14] Sensitivities of LC death certificates of participants and nonparticipants were compared to estimate differential misclassification bias. At study cutoff, the vital status of the PREDICA cohort subjects were ascertained by linkage with the Lombardy Health Registry of all residents in Lombardy Region,[15] which allowed to trace 5730 subjects (98.5% of cohort). For 85 subjects who left their community and were not traceable through the Lombardy Health Registry, investigations were done among persons' next-of-kin and through demographic services, without obtaining proof of their death; therefore, these 85 subjects (1.5% of cohort) were arbitrarily considered alive at the end of study.

To analyze mortality profile, we classified the causes of death by disease groups, according to the International Classification of Diseases, Edition IX (ICD-IX), as follows: LC (ICD-IX: 162.2–162.9); all cancers other than LC (ICD-IX: 140–162.0, 163–239); cardiovascular diseases (ICD-IX: 390–459); respiratory diseases (ICD-IX: 460–519); noncan-

cer disease other than cardiovascular or respiratory (ICD-IX: 001–139, 240–389, 520–999); and all cause.

Statistical Methods

Demographic data and clinical features were reported as frequency, percentage, mean or median, as appropriate; 95% confidence interval (95% CI) and interquartile range were shown. Differences between participants and nonparticipants were tested by Student *t* test, χ^2 test, Fisher's exact test, or Mann-Whitney *U* test as required by type and distribution of variables.

Age-standardized LC incidence in the PREDICA towns general population of smokers was estimated assuming:

annual LC incidence in PREDICA towns in 1997–2006 similar to that recorded in 2000, the year for which official data of LC incidence for Varese Province towns are available;

proportion of LC cases occurring in smokers in PREDICA towns was 86% of all LC cases, as documented in year 2000 in Varese Province[11];

calculation of number of smokers for each age class using gender and 10-year age class percentages of smokers as reported by Istituto Nazionale di Statistica, Italy[16];

LC incidence/1000 person-years adjusted for 2.80 male/female ratio observed in the PREDICA cohort (Table 1);

LC incidence/1000 person-years age standardized by the direct method using as reference the European standard population.[17,18]

The LC incidence in the whole PREDICA cohort was divided by LC incidence in age-matched, gender-matched, time period-matched, smoker status-matched (current versus former smoker), and PREDICA towns location-matched population,[19] yielding the LC incidence standardized rate ratio (SRR). Multivariate logistic regression model was used to analyze the association between screening participation or LC incidence in participants and LC risk factors (familial LC, occupational risk, and pack-years of smoking); odd ratios and 95% CI were estimated.

Age-standardized mortality rates/1000 person-years were calculated by the direct method using as reference the European standard population, the notional standard reference population defined in 1976,[17,18] considering the subjects' age as of the date of study start.

All-cause and disease group-specific MRRs for screening participants and nonparticipants were computed by divid-

TABLE 1. Main Demographic Features and Lung Cancer (LC) Risk Factors of the PREDICA Cohort

Demographic Features and LC Risk Factors	PREDICA Cohort (5815 Subjects)	Participants Subcohort (1244 Subjects; 21% of Cohort)	Nonparticipants Subcohort (4571 Subjects; 79% of Cohort)	*p*
Mean age, yr (95% CI)	56.6 (56.4–56.8)	55.6 (55.2–56.0)	56.8 (56.6–57.1)	<0.001[a]
Age distributions at start of project, by age ranks				<0.0001[b]
45–54 (%)	46.1	50.3	45.0	
55–64 (%)	34.3	36.0	33.9	
65–74 (%)	19.5	13.7	21.1	
Male (%)	73.7	74.6	73.5	0.420[b]
Mean age in males, yr (95% CI)	56.9 (56.7–57.2)	55.9 (55.5–56.4)	57.2 (56.9–57.5)	<0.0001[a]
Mean age in females, yr (95% CI)	55.6 (55.2–56.0)	54.6 (53.8–55.4)	55.8 (55.3–56.3)	0.018[a]
Education level[c]				0.219[b]
Subjects with intermediate school or lower education, n (%)	1697 (80.9)	422 (82.7)	1275 (80.3)	
Subjects with high school or higher education, n (%)	401 (19.1)	88 (17.3)	313 (19.7)	
Smoker status[d]				0.019[b]
Current smoker, n (%)	4208 (76.3)	865 (79.1)	3343 (75.6)	
Former smoker, n (%)	1310 (23.7)	228 (20.9)	1082 (24.4)	
Pack-yr smoked[e] median (IQR)	32.8 (22.8–46.0)	35.4 (26.6–49.0)	31.6 (21.6–45.0)	<0.0001[f]
Subjects with family history of LC[g], n (%)	157 (5.92)	95 (8.95)	62 (3.90)	<0.0001[b]
Subjects with occupational risk of LC[h], n (%)	817 (28.7)	474 (43.5)	343 (19.5)	<0.0001[b]
Person-yr of observation	51 606.9	11 315.1	40 291.8	—
Mean person-yr of observation (95% CI)	8.87 (8.83–8.91)	9.10 (9.02–9.17)	8.81 (8.76–8.87)	<0.0001[a]

Comparison of CXR screening participants vs. nonparticipants. Mean values are presented with 95% confidence intervals (95% CI). The data of pack-yr smoked, family history of lung cancer and occupational risk were all exactly available for 1779/5815 subjects or 30.6% of the PREDICA cohort.

[a] Student *t* test.
[b] χ^2 test.
[c] Available in 2098 subjects (36.1% of cohort).
[d] Available in 5518 subjects (94.9% of cohort).
[e] Available in 3824 subjects (65.8% of cohort).
[f] Mann-Whitney *U* test.
[g] Available in 2650 subjects (45.6% of cohort).
[h] Available in 2845 subjects (48.9% of cohort).
IQR, interquartile range; CXR, chest x-ray.

ing age-standardized mortality rate of participants by age-standardized mortality rate of nonparticipants. The Cox regression model was also used to estimate the hazard risk of death in participants versus nonparticipants.

Poisson distribution was assumed to estimate the difference in LC incidence and in disease group-specific or all-cause mortality rate between participants and nonparticipants, with 95% CI. Differences with $p < 0.05$ were considered significant. Analysis was performed at the Department of Medicine and Public Health, University of Verona (A.Po., W.M.), using Stata software (Stata Corporation, College Station, TX). This study was approved by the Varese Hospital and Health District Ethics Committee.

RESULTS

Participation Rate and Person-Years of Observation

Participants in screening were 1244 of 5815 invited subjects (21%), hereafter referred to as participants subcohort (Table 1). The median number of screening rounds underwent by each subject, including baseline examination, was three instead of the five expected, despite reminder calls. There were 278 other subjects (5% of cohort) who signed consent for screening but did not participate. The remaining 4293 subjects (74% of cohort) did not accept the invitation to participate in screening. Cumulatively 4571 subjects (79% of cohort) did not participate; hereafter, they are referred to as nonparticipants subcohort. In the whole cohort, the cumulative person-years of observation were 51 606.9.

Risk Factors

Demographic and risk features are given in Table 1. The pattern of data missingness is reported in Table 1. The data of three LC risk factors considered for regression analysis (pack-years, family history of LC, and occupational risk) were all exactly present in 1779 of 5815 subjects or 30.6% of the whole PREDICA cohort.

Participants were younger than nonparticipants ($p < 0.001$), and age distributions within subcohorts were different ($p < 0.0001$). The subcohorts displayed similar gender ratio ($p = 0.420$) and education level ($p = 0.219$). Participants featured more pack-years of smoking ($p < 0.0001$) and higher prevalence of current smokers ($p = 0.019$), family history of LC ($p < 0.0001$), and occupational risk ($p < 0.0001$).

Incidence of LC

During the 9.5-year observation period, 245 diagnoses of LC were made (67 participants, 178 nonparticipants). The proportion of histology/cytology-confirmed LC diagnoses was higher in participants than in nonparticipants ($p = 0.028$; Table 2). The proportion of LC diagnoses that were histologically confirmed in nonparticipants was essentially identical to that observed in the year 2000 in the Varese general population (82%).[11] Distribution of histologic types was similar in the subcohorts ($p = 0.573$; Table 2).

The age-standardized LC incidence/1000 person-years in participants was 7.88 (95% CI, 5.52–10.24), significantly higher ($p < 0.001$) than in nonparticipants (4.65 [95% CI, 3.94–5.37]). Age-standardized LC incidence observed in the whole PREDICA cohort was 5.15 (95% CI, 4.47–5.84)/1000 person-years, while the LC incidence expected in the PREDICA towns-matched population was 4.81 (95% CI, 4.06–5.68)/1000 person-years. The PREDICA cohort LC incidence SRR was 1.07 (95% CI, 0.87–1.31).

TABLE 2. Histopathologic Lung Cancer Diagnoses in the PREDICA Cohort

	PREDICA Cohort (5815 Subjects)	Participants Subcohort (1244 Subjects)	Nonparticipants Subcohort (4571 Subjects)	*p*
No. diagnosed lung cancers[a] (% subjects)	245 (4.2)	67 (5.4)	178 (3.9)	
Histologically confirmed cases, n (%)	211 (86)	63 (94)	148 (83)	0.028[b]
Squamous cell carcinoma, n (%)	81 (38)	27 (43)	54 (37)	0.573[c]
Adenocarcinoma, n (%)	69 (33)	18 (29)	51 (34)	
Small cell carcinoma, n (%)	29 (14)	11 (17)	18 (12)	
Undetermined NSCLC, n (%)	32 (15)	7 (11)	25 (17)	

Comparison of CXR screening participants vs. nonparticipants.
[a] Lung cancer: ICD-IX 162.2–162.9.
[b] Fisher's exact test.
[c] χ^2 test.
NSCLC, nonsmall cell lung cancer; CXR, chest x-ray.

Mortality

During the study period, 866 subjects died (118 participants, 748 nonparticipants). Crude mortality rates for all causes, for LC, and for other disease groups considered are presented in Table 3.

Age-standardized MRRs are presented in Table 4. Participants had lower age-standardized all-cause mortality (MRR, 0.67; $p < 0.0001$) and a pattern of lower age-standardized mortality for all-disease groups considered except LC, relative to nonparticipants. In fact, as shown in Table 4, participants had lower mortality for all cancers other than LC (MRR, 0.61; $p = 0.003$), cardiovascular diseases (MRR, 0.38; $p < 0.0001$), and noncancer disease other than cardiovascular or respiratory (MRR, 0.57; $p = 0.019$), relative to nonparticipants. The lower mortality from respiratory diseases of participants was borderline significant (MRR, 0.51; $p = 0.068$). These results suggest that participants self-selected as healthier subjects.

As an alternative method to direct standardization, to control the age-related confounding, we used the Cox regression model; the hazard risk of death in participants versus nonparticipants was 0.66 (95% CI, 0.54–0.80), essentially identical to the 0.67 all-cause MRR obtained by direct age standardization.

During the study period, 179 subjects (40 participants, 139 nonparticipants) died of LC. Because of exclusion from

TABLE 3. Crude Mortality Rates (per 1000 Person-yr) with 95% Confidence Intervals (95% CI) by Causes in the PREDICA Cohort

Cause of Death	PREDICA Cohort (5815 Subjects)	Participants Subcohort (1244 Subjects)	Nonparticipants Subcohort (4571 Subjects)	p^a
Lung cancer (LC)[b]				0.892
n	179	40	139	
Rate	3.47	3.53	3.45	
95% CI	2.97–4.02	2.52–4.81	2.90–4.07	
All cancers other than LC				<0.001
n	269	36	233	
Rate	5.21	3.18	5.78	
95% CI	4.61–5.87	2.23–4.40	5.06–6.58	
Cardiovascular diseases[c]				<0.0001
n	245	22	223	
Rate	4.75	1.94	5.53	
95% CI	4.17–5.38	1.22–2.94	4.83–6.31	
Respiratory diseases[d]				0.048
n	55	6	49	
Rate	1.06	0.53	1.22	
95% CI	0.80–1.39	0.18–1.15	0.90–1.61	
Noncancer disease other than cardiovascular or respiratory				0.008
n	118	14	104	
Rate	2.29	1.24	2.58	
95% CI	1.89–2.74	0.68–2.08	2.11–3.13	
All cause				<0.0001
n	866	118	748	
Rate	16.78	10.43	18.56	
95% CI	15.68–17.94	9.63–12.49	17.26–19.94	

Comparison of CXR screening participants vs. nonparticipants.
[a] χ^2 test.
[b] Lung cancer: ICD-IX 162.2–162.9.
[c] Cardiovascular diseases: ICD-IX 390–459.
[d] Respiratory diseases: ICD-IX 460–519.
CXR, chest x-ray.

enrolment of individuals with ascertained or suspected LC, in the first year of study, only one LC death (participant) was recorded in the cohort. Sensitivity of LC death certificates was high: 164 of 179 (91.6%) LC death certificates were confirmed by the mortality review committee. Sensitivity of LC death certificates was similar in participants and nonparticipants (92.5% versus 91.4%; $p = 0.819$). Importantly, participants had higher LC mortality than nonparticipants (MRR, 1.40; $p = 0.031$).

A group of PREDICA cohort subjects (n = 1779; 30.6% of cohort) with complete information on LC risk factors was available for further analysis. Preliminary evaluation showed that the complete information group (30.6% of cohort) and the incomplete information group (69.4% of cohort) had similar mean age (participants, 55.4 versus 55.7 years, respectively; $p = 0.578$; nonparticipants, 56.7 versus 57.2 years, respectively; $p = 0.112$), gender distribution (participants, 73.8% versus 75.0% male, respectively; $p = 0.628$; nonparticipants, 73.7% versus 72.6% male, respectively; $p = 0.112$) and age-standardized all-cause mortality rate/1000 years (participants, 13.78 versus 15.67, respectively; $p = 0.418$; nonparticipants, 21.21 versus 25.21 respectively; $p = 0.103$). To analyze the association between screening participation or LC diagnosis and LC risk factors, the multivariate logistic regression model was performed only on the complete information group (Table 5).

Participation in screening significantly correlated with familial LC ($p < 0.001$), occupational risk ($p < 0.0001$), and pack-years smoked ($p = 0.002$). Incident LC also correlated with these predictors, although occupational LC risk did not reach statistical significance ($p = 0.089$; Table 5).

DISCUSSION

In the PREDICA LC screening, the participation rate was low (21%), similar to the 20 to 33% observed in Japanese population LC screenings.[20,21] A relevant question is to what extent the low attendance in LC screening and the volunteer effect impact on LC mortality in population-based LC screening. The features of our cohort study, population-based, nonrandomized, and with essentially complete (98.5%) long-term follow-up allow the evaluation of the effects of self-selection on the mortality profile of LC screening participants. In the PREDICA cohort, we observed a volunteer effect character-

TABLE 4. Cumulative Age-Standardized Mortality Rates (per 1000 Person-yr) with 95% Confidence Intervals (95% CI) by Causes in the PREDICA Cohort

Cause of Death	PREDICA Cohort (5815 Subjects)	Participants Subcohort (1244 Subjects)	Nonparticipants Subcohort (4571 Subjects)	MRR	p^a
Lung cancer (LC)[b] (95% CI)	3.71 (3.14–4.23)	4.95 (3.10–6.86)	3.53 (2.93–4.15)	1.40 (1.03–1.91)	0.031
All cancers other than LC (95% CI)	5.23 (4.62–5.95)	3.48 (2.17–4.78)	5.72 (4.95–6.48)	0.61 (0.41–0.91)	0.003
Cardiovascular diseases[c] (95% CI)	5.37 (4.65–6.08)	2.25 (1.11–3.34)	5.99 (5.17–6.81)	0.38 (0.22–0.63)	<0.0001
Respiratory diseases[d] (95% CI)	1.29 (0.93–1.66)	0.71 (0.08–1.42)	1.39 (0.98–1.80)	0.51 (0.24–1.06)	0.068
Noncancer disease other than cardiovascular or respiratory (95% CI)	2.63 (2.13–3.13)	1.63 (0.58–2.68)	2.85 (2.27–3.42)	0.57 (0.34–0.92)	0.019
All cause (95% CI)	18.29 (17.00–19.58)	13.01 (10.14–15.89)	19.48 (18.03–20.94)	0.67 (0.53–0.84)	<0.0001

Comparison of CXR screening participants vs. nonparticipants. Mortality rate ratio (MRR): age-standardized cause-specific mortality rate in participants/age-standardized cause-specific mortality rate in nonparticipants.

[a] p value testing MRR = 1.

[b] Lung cancer: ICD-IX 162.2–162.9.

[c] Cardiovascular diseases: ICD-IX 390–459.

[d] Respiratory diseases: ICD-IX 460–519.

CXR, chest x-ray.

TABLE 5. Predictors of Being Participant in Lung Cancer (LC) Screening and of Incident LC Diagnosis

Outcome	Predictor	OR (95% CI)	p
Participant in LC screening	Familial LC	1.99 (1.31–3.01)	<0.001
	Occupational LC risk	4.42 (3.52–5.54)	<0.0001
	Pack-yr unit increment	1.01 (1.00–1.01)	0.002
Incident LC diagnosis	Familial LC	2.02 (1.11–3.68)	0.021
	Occupational LC risk	1.38 (0.95–1.99)	0.089
	Pack-yr unit increment	1.02 (1.01–1.03)	<0.0001

The analysis was performed in 1779 subjects (30.6% of the PREDICA cohort) for whom all predictors were available. Odds ratios and 95% confidence interval (OR, 95% CI) are shown.

ized by significantly lower all-cause mortality and by significantly lower mortality for several disease groups, except LC. After age standardization, the all-cause mortality in participants remained significantly lower, indicating that self-selection had a healthy effect independent of age.

Participants had significantly higher age-standardized LC incidence ($p < 0.001$) and LC mortality (LC MRR, 1.40; $p = 0.031$) relative to nonparticipants. The higher LC incidence in participants was expected, likely because of higher prevalence of LC risk factors observed and because of some LC overdiagnosis from CXR screening.[22] Only 30.6% of the PREDICA cohort featured complete LC risk information; therefore, the significantly higher familial LC, occupational risk, and pack-years smoked estimated among participants should be cautiously interpreted, although the complete information group and the rest of the cohort were similar for age, gender distribution, and age-standardized all-cause mortality rate. Similarly, the logistic regression analysis, which suggested that the three LC risk factors were predictors of participation in screening and of developing incident LC, should be interpreted cautiously.

The distribution of LC histologic types was similar in the two subcohorts and similar to that of LCs diagnosed in the Varese smokers population in 2000.[11] The 40% higher LC mortality observed in participants is a large difference, for which several possible causes are worthy of consideration. It is unrealistic to postulate a lethal effect of the screening CXRs, an effect that does not occur in clinical practice.

Enrolment of self-selected symptomatic smokers as screenees might cause an overestimation of LC mortality; the excess of LC deaths in participants, however, cannot be justified in this way, because during the first year of study, only one subject died of LC in the participants subcohort. Death certificate errors and sticky-diagnosis bias might also determine overrepresented LC mortality in participants by selective misclassification. Histologic confirmation rate of LCs diagnosed in participants in our study was 11% higher than in nonparticipants, a modest difference that would account for only slight overrepresentation of participants' LC mortality. Sensitivity of LC death certificates was high and similar in participants and nonparticipants (92.5 and 91.4%, respectively), suggesting that selective misclassification did not contribute appreciably to the 1.40 LC MRR recorded. A modest misclassification rate of deaths was observed also in the Mayo LC screening project[14] and in population-based mammographic screening.[23]

Selection bias seems to be the main cause of the mortality difference between participants and nonparticipants in the PREDICA LC screening project. We hypothesize that, in our study, the decision of smokers to attend screening was influenced by their awareness of being at higher LC risk. This hypothesis is supported by a recent study showing that smokers who perceived themselves at higher risk were the most interested in screening.[24] Current smokers and heavy smokers perceived their LC risk as higher than former smokers and light smokers; moreover, smokers with a family history of LC perceived a higher LC risk.[24] Also, in the setting of colorectal screening, a family history of colorectal cancer correlated with increased screening attendance.[25] In the PREDICA cohort, 100% of subjects were aware of their smoking-related

LC risk, illustrated in the screening invitation letter. Moreover, in smokers with a family history of LC or with occupational/environmental exposure risk, increased awareness of LC risk might be prompted also by family doctor, family members, or peers. It is not surprising that a low interest in screening was found in a survey study of light smokers (median: 10 pack-years), with only 21% of the group warned about LC risk.[26]

Our study has limitations. No data on the income level were available for the cohort subjects; we used the education level as surrogate measure of income level. The finding of a slightly lower education level among the participants in our screening is unusual, compared with most screening studies. The difference could be explained by a higher participation in LC screening of subjects with occupational risk and with lower education level, as suggested by the greater proportion of subject with occupational risk in lower versus higher education level that we observed (39.5% versus 20.0%, respectively; $p = 0.001$).

Another limitation of this study is the long duration of enrolment, lasting 4.5 years, a period similar to that of the Mayo Lung Project enrolment.[27] In the real world of cancer screening, when thousands of candidates are involved, enrolment over long time is inevitable, and in our study, it might have slightly inflated all-cause mortality of nonparticipants.

Compliance with annual screening schedule was suboptimal (three CXR examinations per participant), possibly compromising the effectiveness of LC screening; however we did not address the issue of low compliance with CXR exams, because the aim of this report was to estimate self-selection, not to assess screening effectiveness.

Two important questions are whether the PREDICA cohort was complete of all smokers eligible for screening resident in the public GP practices considered and whether it represented the Varese smokers population. As to the first question, the preliminary recruitment list of all smokers aged 45 to 75 years included 5925 subjects, i.e., 20.6% of target age-group residents in the 50 GP practices; this proportion of smokers is essentially identical to the 21% of smokers in the same age-group Lombardy general population, as surveyed in 2000.[16] Moreover, recruited smokers were resident in 44 towns scattered in Varese Province, and the LC incidence SRR of the PREDICA cohort was 1.07, suggesting that the PREDICA cohort LC incidence represented that of Varese smokers.

In conclusion, this population-based study of smokers shows that participation in CXR screening is associated with a healthy effect and a higher LC mortality risk. Moreover, our results suggest that the increased LC mortality in screening participants could be because of self-selection of subjects with higher LC risk. To assess the effectiveness of a CXR screening intervention in smokers on a population level, the possibility of higher LC mortality risk in self-selected participants must be considered. If subjects at risk were overrepresented among screening participants, effectiveness of screening on a population level could theoretically increase; the evaluation of the impact of self-selection on population-based LC screening effectiveness, however, is beyond the purpose of this study and requires further investigations.

ACKNOWLEDGMENTS

Supported by University of Insubria Research Funds FAR1997-FAR2007; Lions Clubs District 108 Ib-1 annual grants 1997–2008. We also acknowledge Lions Club Varese Host; Associazione PREDICA Onlus, Varese; Associazione AMARE, Como; and Associazione P Giancola per la Ricerca sul Cancro, Como, for financial support in the years 1997–2008.

We thank P. Zeli of ASL Varese Health Authority, S. Pisani and D. Bonarrigo of the Epidemiology Observatory of the Varese Province for their advice, and the residents of the Center for Thoracic Surgery, University of Insubria, Varese, for their help in performing the LC screening project. We also thank T. Nakayama and T. Suzuki of the Division of Epidemiology, Osaka Medical Center for Cancer and Cardiovascular Disease for comments on the interpretation of results.

L.D., A.I., and A.Po. were responsible for study development and wrote the report. A.I., M.P., V.D., and A.Pa. performed entry and protection of data and quality assurance monitoring. The mortality review committee responsible for reevaluation of death certificates and final classification of causes of death included clinicians (L.D., A.I., and N.R.), a pathologist (F.S.), an epidemiologist (A.Po.), and a statistician (W.M.).

APPENDIX A

List of 50 General Practitioners of the Italian National Health System working in the Varese Province who collaborated on the PREDICA project: Adreani L, Agosti C, Angelini M, Baj A, Bellingreri V, Borgese S, Bossi P, Brandolino N, Castelli N, Castiglioni D, Cocquio F, Colmegna A, Colombo G, Cordani A, Cova G, Damiani M, Daverio C, Diana R, Dozio F, Gandini T, Giacomino M, Giuliani C, Incremona M, Ligabò M, Lunardon M, Magnaghi F, Marangoni L, Marcianò M, Mazza G, Morenghi R, Nicora F, Pala R, Pasquali C, Pisani S, Pisciotta M, Pizzi M, Rizzuto G, Roberto V, Rossi N, Sassi G, Scienza G, Sinapi D, Tamborini M, Testorelli M, Tonello E, Trotta L, Valente A, Vigoni V, Zanzi C, and Zecchini D.

APPENDIX B

List of "PREDICA towns" in the Varese Province: Albizzate, Arcisate, Azzate, Besano, Besozzo, Bisuschio, Bodio Lomnago, Brezzo di Bedero, Brissago Valtravaglia, Brunello, Buguggiate, Cantello, Caronno Varesino, Cassano Magnano, Cittiglio, Clivio, Cuasso al Monte, Cugliate Fabiasco, Cunardo, Cuveglio, Daverio, Gallarate, Galliate Lombardo, Gavirate, Gazzada Schianno, Germignaga, Induno Olona, Laveno Mombello, Lonate Pozzolo, Luino, Malnate, Marchirolo, Mesenzana, Morazzone, Mornago, Oggiona con Santo Stefano, Porto Ceresio, Rancio Valcuvia, Saltrio, Samarate, Solbiate Arno, Valganna, Varese, and Viggiù.

REFERENCES

1. Richardson A, Wells JE. Breast cancer screening: the effect of self selection for screening on comparisons of randomized controlled trials. *J Med Screen* 1997;4:16–18.
2. Ciatto S, Gervasi G, Gorini G, et al. Prostate cancer specific mortality in

the Florence screening pilot study cohort 1992–1993. *Eur J Cancer* 2006;42:1858–1862.
3. Blom J, Yin L, Lidén A, et al. A 9-year follow-up study of participants and nonparticipants in sigmoidoscopy screening: importance of self-selection. *Cancer Epidemiol Biomarkers Prev* 2008;17:1163–1168.
4. Pinsky PF, Miller A, Kramer BS, et al. Evidence of a healthy volunteer effect in the prostate, lung, colorectal and ovarian cancer screening trial. *Am J Epidemiol* 2007;165:874–881.
5. Paci E, Alexander FE. Study design of randomized controlled clinical trials of breast cancer screening. *J Natl Cancer Inst Monogr* 1997;22: 21–25.
6. Sagawa M, Tsubono Y, Saito Y, et al. A case-control study for evaluating the efficacy of mass screening program for lung cancer in Miyagi Prefecture, Japan. *Cancer* 2001;92:588–594.
7. Sobue T. A case-control study for evaluating lung cancer screening in Japan. *Cancer* 2000;89S:2392–2396.
8. Dominioni L, Imperatori A, Rovera F, et al. Lung cancer screening in cigarette smokers in the province of Varese, Italy. *Cancer* 2000;89S: 2345–2348.
9. Matakidou A, Eisen T, Houlston RS. Systematic review of the relationship between family history and lung cancer risk. *Br J Cancer* 2005;93: 825–833.
10. International Agency for Research on Cancer. *Wood Dust and Formaldehyde. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*, Vol. 62. Lyon, France: IARC, 1997.
11. Imperatori A, Harrison RN, Leitch DN, et al. Lung cancer in Teesside (UK) and Varese (Italy): a comparison of management and survival. *Thorax* 2006;61:232–239.
12. Associazione Italiana Registri Tumori. Varese Cancer Registry. Available from: http://www.registri-tumori.it/. Accessed December 17, 2008.
13. Azienda Sanitaria Locale della Provincia di Varese. Varese Province Epidemiology Observatory. Available from: http://www.asl.varese.it/. Accessed December 17, 2008.
14. Doria-Rose VP, Marcus PM. Death certificate provide an adequate source of cause of death information when evaluating lung cancer mortality: an example from the Mayo Lung Project. *Lung Cancer* 2009;63:295–300.
15. Regione Lombardia. Servizi Informativi Socio-Sanitari. Available from: https://was.cgi.crs.lombardia.it/SebsGate/. Accessed December 23, 2008.
16. Istat.it. Lifestyle and Health Conditions Year 2000. Available from: http://www.istat.it/dati/catalogo/20010323_01/volume.pdf. Accessed December 23, 2008.
17. Waterhouse JAH, Muier CS, Correa P, et al. (Eds). *Cancer Incidence in Five Continents*, Vol. 3. Lyon, France: IARC, 1976. Pp. 456.
18. European Union Public Health Information System (EUPHIX)—European Standard Population. Available from: http://www.euphix.org/object_document/o5338n27620.htm. Accessed December 23, 2008.
19. Istituto Nazionale di Statistica. Italian Population Census 2001. Resident Population of Varese Province. Available from: http://dawinci.istat.it/daWinci/jsp/MD/excel/g1UH07M07RE4000000.xls. Accessed December 23, 2008.
20. Nishii K, Ueoka H, Kiura K, et al. A case-control study of lung cancer screening in Okayama Prefecture, Japan. *Lung Cancer* 2001;34:325–332.
21. Okamoto N, Suzuki T, Hasegawa H, et al. Evaluation of a clinic-based screening program for lung cancer with a case-control design in Kanagawa, Japan. *Lung Cancer* 1999;25:77–85.
22. Yankelevitz DF, Kostis WJ, Henschke CI, et al. Overdiagnosis in chest radiographic screening for lung carcinoma: frequency. *Cancer* 2003;97: 1271–1275.
23. Goldoni CA, Bonora K, Ciatto S, et al. Misclassification of breast cancer as cause of death in a service screening area. *Cancer Causes Control* 2009;20:533–538.
24. Hahn EJ, Rayens MK, Hopenhayn C, et al. Perceived risk and interest in screening for lung cancer among current and former smokers. *Res Nurs Health* 2006;29:359–370.
25. Blom J, Yin L, Lidén A, et al. Toward understanding nonparticipation in sigmoidoscopy screening for colorectal cancer. *Int J Cancer* 2008;122: 1618–1623.
26. Silvestri GA, Nietert PJ, Zoller J, et al. Attitudes towards screening for lung cancer among smokers and their non-smoking counterparts. *Thorax* 2007;62:126–130.
27. Fontana RS, Sanderson DR, Woolner LB, et al. Lung cancer screening: the Mayo program. *J Occup Med* 1986;28:746–750.

