View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Archivio della Ricerca - Università di Pisa

Multimodal lung cancer screening using the ITALUNG Biomarker Panel and Low Dose Computed Tomography. Results of the ITALUNG biomarker study

Francesca Maria Carozzi[1], Simonetta Bisanzi[1], Laura Carrozzi[2], Fabio Falaschi[3], Andrea Lopes Pegna[4], Mario Mascalchi[5], Giulia Picozzi[6], Marco Peluso[1,*], Cristina Sani[1], Luana Greco[7], Cristina Ocello [8], Eugenio Paci [9] and the ITALUNG Working Group.

[1] Regional Cancer Prevention Laboratory, ISPO-Cancer Prevention and Research Institute, Florence, Italy
[2] Cardiopulmonary Department, University Hospital of Pisa, Pisa, Italy
[3] Radiology Department, University Hospital of Pisa, Pisa Italy
[4] Retired, Pneumonology Department, Careggi Hospital, Florence, Italy
[5] Radiodiagnostic Section, Department of Experimental and Clinical Biomedical Sciences, University of Florence, Florence, Italy
[6] Radiodiagnostic Unit, ISPO-Cancer Prevention and Research Institute, Florence, Italy
[7] Biologist, formerly at ISPO-Cancer Prevention and Research Institute, Florence, Italy.
[8] Statistician, Florence, Italy.
[9] Retired, formerly at ISPO-Cancer Prevention and Research Institute, Florence, Italy.

ITALUNG Working Group members: Francesca Maria Carozzi, BSc, Cristina Maddau, BSc, Simonetta Bisanzi, BSc, Regional Cancer Prevention Laboratory, ISPO-Cancer Prevention and Research Institute, Florence, Italy; Eugenio Paci, MD (PI), Donella Puliti, MSc, Marco Zappa, MD, Cristina Ocello, BSc, PhD, Carmen Visioli, MD, Giovanna Cordopatri, BSc, Francesco Giusti, BSc, PhD, Ida Esposito, BSc, Gianfranco Manneschi, BSc, Unit of Clinical Epidemiology, ISPO-Cancer Prevention and Research Institute, Florence, Italy; Andrea Lopes Pegna, MD, Roberto Bianchi, MD, Cristina Ronchi, MD, Pneumonology Department, Careggi Hospital, Florence, Italy; Laura Carrozzi, MD, Ferruccio Aquilini, BSc, Stella Cini, MD, Mariella De Santis, MD, Francesco Pistelli, MD, Filomena Baliva, MD, Antonio Chella, MD, Laura Tavanti, MD, Cardiopulmonary Department, University Hospital of Pisa, Pisa, Italy; Michela Grazzini, MD, Florio Innocenti, MD, Ilaria Natali, BSc, Pneumonology Department, Hospital of Pistoia, Pistoia, Italy; Mario Mascalchi, MD, PhD, Maurizio Bartolucci, MD, Elena Crisci, MD, Agostino De Francisci, MD, Massimo Falchini, MD, Silvia Gabbrielli, MD, Giulia Picozzi, MD, Giuliana Roselli, MD, Andrea Masi, MD, Radiology Department, Careggi Hospital, University of Florence, Florence, Italy; Fabio Falaschi, MD, Luigi Battola, MD, Anna Lisa De Liperi, MD, and Cheti Spinelli, MD, Radiology Department, University Hospital of Pisa, Pisa, Italy; Letizia Vannucchi, MD, Alessia Petruzzelli, MD, Davide Gadda, MD, Anna Talina Neri, MD, and Franco Niccolai, MD, Radiology Department, Hospital of Pistoia, Pistoia, Italy; Luca Vaggelli, MD, Nuclear Medicine Department, Careggi Hospital, Florence, Italy; Alessandra Vella, MD, PhD, Nuclear Medicine Department, Le Scotte University Hospital, Siena, Italy; Alberto Janni, MD, Thoracic Surgery Department, Careggi Hospital, Florence, Italy; Alfredo Mussi, MD and Marco Lucchi, MD, Thoracic Surgery Department, University Hospital of Pisa, Pisa, Italy; Camilla Comin, MD: Pathology Department, Careggi Hospital, University of Florence, Florence, Italy; Gabriella Fontanini, MD, Adele Renza Tognetti, MD, Pathology Department, University Hospital of Pisa, Pisa, Italy.

Short title: ITALUNG biomarker panel study.

Keywords: Lung cancer screening, sputum, peripheral blood, plasma DNA, genomic instability.

Abbreviations: LDCT: low-dose computed tomography; NLST: national lung cancer screening randomized trial; IBP: ITALUNG biomarker panel; RCT: randomised clinical trials; NCN: non calcific nodule; MSI: microsatellite instability; LOH: loss of heterozygosity.

Article category: Cancer epidemiology.

Conflict of Interest: The authors declare no conflict of interest.

Grant sponsor: Supported by a grant (n. 40/50) from the Italian League for the Fight against Cancer (LILT) and the Italian Ministry of Health (Progetto Integrato di Oncologia, area tematica 1, sperimentazione di modelli di intervento di prevenzione primaria e secondaria) for molecular analysis. The Italung RCT was entirely funded by the Tuscany Region Department.

*** Corresponding author:** Marco Peluso, Regional Cancer Prevention Laboratory, ISPO-Cancer Prevention and Research Institute, 50139, Florence.

Novelty and Impact: LDCT screening is a complex process with high probability of recall for assessment of Non Calcific Nodules. Biomarkers are possible contributors at different stages of LDCT lung cancer Screening. The ITALUNG Biomarker Panel (IBP) combined plasma DNA quantification (cut-off 5ng/ml)

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as an 'Accepted Article', doi: 10.1002/ijc.30727

This article is protected by copyright. All rights reserved.

Accepted Article

and loss of heterozygosity and microsatellite instability, showed high sensitivity for lung cancer detection at baseline test. Use of IBP as multimodal screening could contribute to reduce the burden of lung cancer screening.

Accepted Article

John Wiley & Sons, Inc.

This article is protected by copyright. All rights reserved.

Accepted Article

Abstract

Asymptomatic high-risk subjects, randomized in the intervention arm of the ITALUNG trial (1406 screened for lung cancer), were enrolled for the ITALUNG biomarker study (n = 1356), in which samples of blood and sputum were analysed for plasma DNA quantification (cut off 5ng/ml), loss of heterozygosity and microsatellite instability. The ITALUNG biomarker panel (IBP) was considered positive if at least one of the two biomarkers included in the panel was positive. Subjects with and without lung cancer diagnosis at the end of the screening cycle with LDCT (n = 517) were evaluated. Out of 18 baseline screen detected lung cancer cases, 17 were IBP positive (94%). Repeat screen-detected lung cancer cases were 18 and 12 of them positive at baseline IBP test (66%). Interval cancer cases (2-years) and biomarker tests after a suspect Non Calcific Nodule follow-up were investigated. The single test versus multimodal screening measures of accuracy were compared in a simulation within the screened ITALUNG intervention arm, considering screen-detected and interval cancer cases. Sensitivity was 90% at baseline screening. Specificity was 71%% and 61% for LDCT and IBP as baseline single test, and improved at 89% with multimodal, combined screening. The positive predictive value was 4.3% for LDCT at baseline and 10.6% for multimodal screening. Multimodal screening could improve the screening efficiency at baseline and strategies for future implementation are discussed. If IBP was used as primary screening test, the LDCT burden might decrease of about 60%.

This article is protected by copyright. All rights reserved.

Accepted Article

Introduction

Lung cancer is currently the leading cause of cancer death in the world. Although the rates of lung cancer mortality have started to decrease in most industrialized countries, the long time lag between the peak of cigarette consumption and lung cancer development will assure a long life for the epidemic. Smoking attributable deaths are also projected to increase due to the surge in tobacco consumption among young people and in developing countries [1, 2]. Additionally, at the time of diagnosis, lung cancer is often in an advanced stage of disease, with a 5-year survival lower than 20% (https://seer.cancer.gov/statfacts/html/lungb.html). The US National Lung Cancer Screening Trial (NLST), a randomized trial utilizing LDCT, showed a significant reduction in the overall and lung cancer specific mortalities when compared to screening with chest radiography [3]. The ITALUNG RCT is one of the European randomised trials to assess the benefit-to-harm ratio of screening with LDCT [4].

Biomarkers have been considered possible contributors at different stages of the lung cancer screening process [5]. The use of biomarkers for predicting the occurrence of the disease is one of the major aims of recent research but, as yet, there have been no conclusive results. The better identification of high-risk subjects using biomarkers or the integration of biomarker tests with LDCT in a multimodal screening has received less attention. In this paper the level of plasma DNA and genomic instability (MSI/LOH) in both sputum and plasma DNA are evaluated for their potential as identifiers of subjects in the phase of field cancerisation or as supplementary screening tests [6]. The optimisation of both sensitivity and specificity by means of the combined use of biomarkers and LDCT was assessed in sputum and plasma samples provided by ITALUNG subjects.

Materials and Methods

Study design

The ITALUNG study for the early detection of lung cancer using LDCT is a randomised screening trial based on the selection of individuals aged 55 to 69 years who are heavy smokers or ex-smokers with a smoking exposure history of at least 20 pack-years in the last 10 years. The study is registered at Clinical Trial.gov (ID = NCT02777996). The study design and characteristics of the ITALUNG subjects have been presented [7]. In brief, letters were sent to residents in three districts of the region of Tuscany. Subjects were randomised to the screening intervention or the usual care (no screening) arms. The controls were followed

This article is protected by copyright. All rights reserved.

Accepted Article

up in usual care. The intervention arm (n = 1613) was invited to attend a screening cycle consisting of a baseline and 3 repeat screening rounds at 1-year intervals. 1406 asymptomatic high-risk subjects were screened at baseline and participated in the screening cycle [8]. In the presence of Non Calcific Nodule (NCN) ≥ 5mm at baseline or ≥ 3 mm at a repeat round, the study provided recall for further assessment with diagnostic work-up, in accordance with a strict protocol.

The ITALUNG biomarker study was approved by the Ethical Committees of Florence (n. 23/2003/CEL) and each participating centre. Out of 1406 screened, 1356 subjects gave their individual consent to give a sample of blood and sputum at baseline LDCT screening for storage in the ITALUNG biobank. A second sample of blood and sputum was requested when subjects were recalled for further assessment at baseline or repeat rounds. Our analysis, which covered the entire screening cycle, included those subjects diagnosed with lung cancer (n = 36 out of 38 screen-detected lung cancer cases in the ITALUNG intervention arm) and a random selection of subjects without lung cancer diagnosis (n = 481 / 1320=36%)), for a total of 517 subjects.

Those subjects without lung cancer diagnosis at the end of the screening cycle were classified into one of 3 LDCT screening groups: (a) the S1 group included subjects defined as “always negative for NCN”, i.e. without lung cancer at the end of the entire screening cycle and never recalled for the assessment of NCN either at baseline LDCT or at annual repeat LDCT (n = 235); (b) the S2 group included subjects defined as “baseline negative, positive for NCN at an annual repeat screening”, i.e. without lung cancer at the end of the entire screening cycle, with a negative result at LDCT baseline, and with an NCN ≥3mm detected at annual repeat LDCT (n = 118); (c) the S3 group included subjects defined as “baseline positive for NCN”, i.e. without lung cancer at the end of the entire screening cycle, but with an NCN ≥5 mm at baseline LDCT, with or without an NCN ≥3mm at annual repeat LDCT (n = 128). All screen-detected lung cancer subjects with a baseline sample available were included and classified according to detection at baseline or at an annual follow-up: (a) the S4 group included subjects with an NCN at baseline LDCT and diagnosed as screen-detected lung cancer after diagnostic work-up (n = 18, lung cancer at baseline test); (b) the S5 group included subjects diagnosed as screen-detected lung cancer at annual repeat LDCT (n = 18, lung cancer at annual repeat test).

This article is protected by copyright. All rights reserved.

Even though the LDCT screening interval was 1 year, interval cancer cases were defined (following a suggestion in the NELSON RCT evaluation [9]) as those clinical cases which, at the end of the ITALUNG RCT follow-up (8.5 years) and based on data from the Cancer Registry of Tuscany, had occurred within 2 years of a negative screening test .[10] There were 6 interval cancers, only 2 of which were after the baseline screening round.

Clinical samples

Samples of peripheral blood and sputum collected at home for 3 consecutive mornings were obtained from each subject and duly processed according to the study protocol. Sputum was collected in 15 cc of CytoLyt Solution (Hologic) and peripheral blood was collected in K3-EDTA, as previously reported [6].

Plasma DNA

The amount of DNA in plasma was determined through the use of Real Time quantitative PCR with amplification of the gene of interest, hTERT, mapped in single copy in the region 5p15.33, according to the protocol already described [6]. The data were analysed by the Real Time PCR 7500 Sequence Detection System (Life Technologies). In the previous study, in which we used a ROC analysis of continuous data, we set the cut-off point to 5 ng DNA/ml of plasma [6]. We used the same cut-off in this study.

Genomic

The analyses of microsatellite instability (MSI, allele shift) and the loss of heterozygosity (LOH) were carried out by studying the alteration of microsatellites relating to 12 loci in 5 different chromosomes: 3p14.2 (D3S1300), 3p21-p23 (D3S1289), 3p26.1 (D3S1263), 3p13 (D3S1566), 5q15 (D5S644), 5q22.2 (D5S2084), 9p22-p23 (D9S157), 9p21 (D9S161), 13q12.3 (D13S171), 13q14 (D13S153), 13q14.1 (D13S263) and 17p13.2 (D17S938) [6].

A sputum sample was scored positive for LOH when a reduction for one allele's intensity of at least 23% was found whereas the required reduction for a plasma sample was set to at least 30%. The greater reduction for plasma provided a cut-off level that avoided false positive results (due to the lower amounts of plasma DNA) and enhanced reproducibility. All clinical samples were assayed twice but when inconsistencies in values around the boundaries were observed a third assay was performed. The presence of microsatellite instability was indicated by the presence of additional series of peaks before and after the peak expected for the markers. A subject was scored positive for MSI/LOH whichever sample was positive.

John Wiley & Sons, Inc.

This article is protected by copyright. All rights reserved.

Accepted Article

Accepted Article

The ITALUNG biomarker panel (IBP) was calculated from the individual MSI/LOH and DNA plasma values (cut-off 5 ng/ml). The IBP was positive if at least one of the two panel biomarkers was positive.

We simulated lung cancer screening strategies based on the LDCT results and analyses of the clinical samples. The proportion of IBP positive subjects was applied to the screened ITALUNG subjects according to the presence or absence of LDCT recall for NCN.

Results

Table 1 shows the characteristics of the 517 subjects included in the biomarker analysis. In total, 68.1% were current smokers at baseline screening, the mean tobacco consumption for all eligible subjects (smokers and ex-smokers) was 43.7 pack years and the mean age 61.1 years.

Table 2 shows biomarker positivity by LDCT screening group - the MSI/LOH and plasma DNA results. The plasma DNA estimate was missing in 13.5% of subjects without lung cancer (n = 65/481) and, also, in one of the annual repeat screen-detected lung cancer cases (n = 1/18). Plasma DNA result was positive in 13.3% of subjects without lung cancer (i.e. groups S1, S2 & S3) The MSI/LOH sputum and blood samples (at 23% and 30% thresholds, respectively) were positive in 34.9% of subjects without lung cancer (i.e. groups S1, S2 & S3) (Fisher's exact test = 0.672). Some 88.9% and 57.1%, respectively, of the baseline and annual repeat test, lung cancer cases were MSI/LOH positive (Fisher Exact test = 0.228).

The IBP was considered positive at baseline test in 39% of the subjects without lung cancer at the end of the screening process (i.e. groups S1, S2 & S3). The number of baseline screen-detected lung cancer cases which were positive at IBP was 17 (94.4%), out of the 18 lung cancers in this group. Twelve of the 18 lung cancers detected at annual repeat LDCT were positive at IBP (66.7%).

At baseline, 517 sampled subjects had been screened (Table 3), from which 146 were LDCT positive, i.e. had had a suspicious finding implying a more intensive workup. Considering baseline screen detected cases only, specificity of LDCT alone (Sp_{LDCT}) was estimated as (371)/(517-18) = 74%. For IBP alone, Sp_{IBP} was = (296-1)/499 = 59%. However, if LDCT and IBP were combined, i.e. considering "screening positive" the subjects with both tests positive and needed of further workup, the Sp_{comb} was = (80+156+214)/499 = 90%. The Positive Predictive Value (PPV) of LDCT and IBP combined (PPV_{Comb}) was = 17/65 = 26%.

John Wiley & Sons, Inc.

This article is protected by copyright. All rights reserved.

Interval cancers after baseline LDCT screening (both negative at baseline LDCT) were 1 positive and 1 negative at baseline IBP. Out of 4 interval cancer cases after a repeated LDCT test (1 to 3 rounds), one was not included in this study, because the subject did not consent to participate in the biomarker study. Others were all negative at baseline LDCT screening test, 1 of them positive and 2 negative at baseline IBP.

Because these results were based on "enriched data", i.e. including all lung cancer cases, but only a selection of screening negative subjects and did not included interval cancer cases, we decided to extrapolate them to a hypothetical baseline screening situation for the entire ITALUNG screened cohort of 1406 subjects to avoid any selection bias. In the ITALUNG study, 420 subjects were recalled at baseline LDCT screening in accordance with the radiological LDCT protocol (30%), 20 lung cancer cases were screen-detected (Detection rate = 1.42%) and 2 interval cancer cases (2-years) were diagnosed [8].

In the simulation, the proportion of IBP results expected in each cell on the basis of the biomarker study was applied to the positive and negative LDCT subjects (Table 4). In Table 5, screen detected and interval (2-years) lung cancer cases observed in the whole screening cycle and participant in the biomarker study are presented by LDCT and IBP result at baseline screening. In Table 6, measures of accuracy at single or multimodal, combined screening are presented. LDCT and IBP as single tests at baseline had equal sensitivity (90%), but not specificity, which was lower for IBP (71% vs 61%), with similar Positive Predictive Values (PPV) (4.3 % vs 3.3%). The simulation of the LDCT and IBP combined, i.e. a multimodal strategy whereby the baseline test result would be used to select subjects for screening, improved screening Specificity (Sp_{Comb} = 89%), as the PPV_{Comb} (10.6%); sensitivity was unchanged (90%).

In 986 LDCT screened subjects negative at baseline (Table 4) , 18 lung cancer cases were screen detected at three repeated rounds and 4 interval cancers (2-years) were diagnosed (2.3%). IBP positive at baseline were estimated as 378 (38.3%). Out of them, 12 cases had been screen detected at repeated LDCT tests and one interval cancer. The probability to be diagnosed with a lung cancer in this LDCT negative, IBP positive group at baseline (13/22=59% of the lung cancer cases) , over the whole screening cycle, was 3.4%.

Each subject recalled for follow-up or for further assessment, following detection of a Non Calcific Nodule at baseline or repeat screening test, was requested to give a further sputum and blood sample for storage in the ITALUNG biobank. Out of the 18 subjects with baseline screen-detected lung cancer, 17 were baseline IBP positive (see Additional material online). Out of them, according to protocol, 7 were followed

This article is protected by copyright. All rights reserved.

Accepted Article

up because of the detection of an NCN and all were IBP positive in the follow-up sample analysis. Out of the 18 lung cancer cases screen-detected at repeat screening, 6 were IBP negative at the baseline IBP test, and at the follow-up test 3 of them switched to an IBP positive result, 2 were confirmed as negative and 1 did not receive the follow-up test. On the contrary, 12 of the 18 screen-detected lung cancers at repeat test were evaluated as IBP positive at baseline LDCT screening test, and 9 of them were confirmed IBP positive at the follow-up IBP test and 3 were not tested.

Discussion

Lung cancer screening with LDCT is a complex care system with high costs and side effects, such as the high rates of recall for further assessment. Since the publication of the NLST results [3], lung cancer screening with LDCT has been implemented in the USA as an effective tool for reducing lung cancer mortality [11]. The offer of the test, according to the most influential guidelines, should be conditional on the level of risk and, further, strict screening protocols should be adopted. There are, however, concerns about the costs and possible side effects [12]. Subsequent evaluations of the NLST have confirmed that the effect of screening could be better achieved by targeting high-risk populations, a valid approach for the prevention of the greatest number of deaths. [13] In a re-analysis of the NLST, the number of lung cancer deaths prevented was shown to increase according to the individual risk quintile at baseline. According to Kovalchik et al. [14] in the NLST, 60% of the participants at high-risk for lung cancer death accounted for 88% of the screening-prevented lung cancer deaths.

The recently published data of the UKLS pilot trial reported a level for the first-round detection rate which was 19% more than in ITALUNG (1.7% vs 1.4%), with the UKLS subjects having a higher risk profile but a lower proportion of current smokers. [15,16] On the basis of evaluation of the existing risk models, the question of who should be screened continues to be debated [13]. The utility, as tools, of hyper-methylation, loss of heterozygosis and plasma DNA, together with the evaluation of microRNA, has been suggested as complementary tests for screening for lung cancer. DNA hyper-methylation is present in many types of tumour and occurs very early in the cancerization phase. Indicators of methylation have been proposed in several studies, but none are yet included in clinical practice [17,18]. In the NELSON trial [19] with a screening population at high risk for lung cancer, DNA hyper-methylation analysis in sputum was assessed as possibly

This article is protected by copyright. All rights reserved.

playing a role in the detection of preclinical disease, but complementary diagnostic markers are needed to improve sensitivity. Several studies have also evaluated the microRNA contribution for their utility as predictors of the growth and aggressiveness of lung cancer. The microRNA signature classifier (MSC) has been, retrospectively, evaluated by Sozzi et al. [20] showing high sensitivity (87%) and high specificity (81%) for cancer. In the COSMOS study sensitivity and specificity for cancer of a different signature were 77.8% and 74.8%, respectively [21,22].

In this analysis, we have confirmed that genetic MSI/LOHs are a sensitive marker of signature of field cancerisation of lung tissue. Lung cancer studies have confirmed that there are highly altered chromosomal regions, and there is evidence that alterations are strongly influenced by smoking habits [23-26] and that they are not just specific for lung cancer. In our study a high MSI /LOH sensitivity for screen-detected lung cancer cases at the baseline LDCT round has been shown, although with moderate specificity. In contrast, plasma DNA (>5ng/ml) has shown higher specificity for lung cancer at baseline screening, but with lower sensitivity.

Considering the strong heterogeneity of lung cancer and the progressive accumulation of genetic and molecular alterations occurring during the process of carcinogenesis, it is possible to achieve better results by evaluating a panel of molecular alterations using more than a single biomarker. We proposed the IBP as such a heterogeneous panel, encompassing different molecular alterations. The IBP, the biomarker panel of MSI/LOH and/or plasma DNA combined results, was tested for accuracy, and has demonstrated good performance for the identification of lung cancers at baseline screening, with a very high sensitivity.

This finding suggests a rapid increase of the DNA Plasma quantitative measure and a long duration for the field cancerisation phase. Nevertheless, the lower sensitivity of biomarkers at repeat tests can also be related to limitations in the reading process at the baseline LDCT screening. A detailed reconsideration of the LDCT reading process at initial screening has been performed in ITALUNG [27] and an evaluation of the characteristics of interval and screen-detected cases jointly considering tumour morphology, radiologic imaging and modality of detection is on-going. An improved understanding of the biomolecular characteristics is indeed needed to confirm the indication of the MSI/LOH loci involved, which were selected *a priori* as being more related to the detection of lung cancer. In this study we noted the association of the positive result at some MSI/LOHs loci, with the loci 1 to 5 (3p14.2, 3p21-p23, 3p26.1, 3p13, 5q15)

John Wiley & Sons, Inc.

This article is protected by copyright. All rights reserved.

Accepted Article

Accepted Article

and 7 to 9 (9p22-p23, 9p21, 13q12.3), using a 30% cut off. A larger prospective study is needed to evaluate the hypothesis of using the most specific MSI/LOH loci.

The diagnostic accuracy measures of LDCT screening were estimated in other RCTs. The sensitivity and specificity of LDCT were measured in the NLST [4] (83.8% and 73.4% at baseline) and also considering screen-detected lung cancer cases over 3 rounds [28]. In the NELSON (84.6% and 98.6%) screen-detected cancers over 3 rounds and interval cancers (2-years)) were included [9]. In the ITALUNG study, there were 1406 active subjects in the intervention arm at baseline who experienced 5293 negative LDCT screening episodes in 4 rounds [8]. The accuracy measures, based on the 4 rounds, were 86.4% (95% C.I.: 70.5%-93.5%) and 80.3% (95% C.I.: 79.9%-82.0%) for sensitivity and specificity, respectively. The Positive Predictive Value of NLST was very comparable to the observed in ITALUNG study, whereas much higher in NELSON study.

Next, we analysed an hypothetical baseline screening situation for the entire ITALUNG screened cohort (Table 6), rather than on only a selection of screening negative subjects which did not include interval cancer cases (Table 3). This choice was based on the attempt to avoid any potential distortion of the estimates due to selection bias. The strength of the extrapolation of the sample biomarker data to the subjects enrolled in the ITALUNG screened arm is the estimate of accuracy measures and follow-up data over the entire lung cancer screening cycle (4 screening rounds), with the inclusion of interval cancers (2-years) as false negative tests for LDCT. In the simulation of the performance of the IBP test, we assumed the interval cancers as true positive if the IBP baseline test result was positive. On the basis of these estimates (Table 6), the IBP as single screening test had comparable sensitivity but much lower specificity of the LDCT (71% vs 61.2%), too low for a single screening test and near to a random selection of subjects. However, sensitivity of the multimodal, combined baseline screening was comparable to single test screening (both LDCT and IBP), but with better specificity and PPV.

Two strategies of multimodal screening were considered. First, LDCT screening as primary test can be offered to all eligible subjects and the combined IBP positive result for LDCT positive subjects will discriminate subjects at higher risk of lung cancer, i.e. improve the decision of further workup, increasing the positive predictive value. In the second strategy, IBP screening is used as primary test. If the result of LDCT is positive, then specificity and PPV are improved.

This article is protected by copyright. All rights reserved.

Accepted Article

In LDCT negative subjects at baseline (Table 3 and Additional material online), baseline IBP is a predictor of a proportion of screen-detected (n = 12 out of 18) and interval lung cancers (n = 1 out of 3) at repeat screening test. The prediction of future occurrence at baseline screening is limited. Understanding the interrelationships between biomarkers and LDCT radiological findings might elevate our ability of prediction, but a follow-up protocol considering IBP combined with LDCT could improve the screening performance at repeated screening. A limitation of this study is the availability of the IBP test for all subjects only at baseline screening. However, subjects recalled for an NCN at repeat screening test were requested to provide a further sputum and blood sample. In most lung cancer cases, the status of the IBP switched to positive in the samples collected near the time of the diagnosis of lung cancer. Also. numbers are small and further studies would be needed to confirm this behaviour.

LDCT screening has relevant costs, conditioned by local conditions and management. Testing MSI/LOH and Plasma DNA (i.e. IBP) is still done in a research setting and an evaluation of costs as a routine practice is under way. In the "LDCT-first strategy", only the LDCT positive subjects would be screened by IBP, according with the screening protocol and information used for the management of NCN recalled for further assessment or follow-up. If we follow the 'believe the negative' rule as Shaw et al. have suggested [29], all subjects who were IBP negative would not be further investigated but would have a repeated biomarker test (and a LDCT if positive) after a screening interval. In the "IBP-first strategy", only subjects with biomarkers positive at baseline would receive LDCT screening and all subjects repeat the test (and LDCT if positive at repetition). The consequence of this stragegy is a potential important reduction of the number of LDCT tests at baseline, about the 40% of the 1406 subjects (843 LDCT tests).

Prior to the translation of this approach into screening practice, the performance, protocol and diagnostic accuracy of multimodal screening should be tested in other lung cancer screening trials which have a comparable biobank of sputum and blood samples and, further, in a new prospective study of high risk subjects selected for LDCT by means of the ITALUNG biomarker panel.

Conflict of interest

The authors declare that they have no financial disclosure or conflict of interest.

Acknowledge

We thank the reviewers for the constructive and thoughtful comments we received.

This article is protected by copyright. All rights reserved.

Accepted Article

References

1. Leon ME, Peruga A, McNeill A, Kralikova E, Guha N, Minozzi S, Espina C, Schüz J. European Code against Cancer, 4th Edition: Tobacco and cancer. *Cancer Epidemiol Biomarkers Prev* 2015;**39** S20-33.

2. IARC. Cancer incidence in five continents. Volume X. *IARC Sci Publ* 2013.

3. National Lung Screening Trial Research T, Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, Fagerstrom RM, Gareen IF, Gatsonis C, Marcus PM, Sicks JD. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med* 2011;**365**: 395-409.

4. Field JK, van Klaveren R, Pedersen JH, Pastorino U, Paci E, Becker N, Infante M, Oudkerk M, de Koning HJ, European Randomized Screening Trial G. European randomized lung cancer screening trials: Post NLST. *Journal of surgical oncology* 2013;**108**: 280-6.

5. Atwater T, Massion PP. Biomarkers of risk to develop lung cancer in the new screening era. *Ann Transl Med* 2016;**4**: 158.

6. Carozzi FM, Bisanzi S, Falini P, Sani C, Venturini G, Lopes Pegna A, Bianchi R, Ronchi C, Picozzi G, Mascalchi M, Carrozzi L, Baliva F, et al. Molecular profile in body fluids in subjects enrolled in a randomised trial for lung cancer screening: Perspectives of integrated strategies for early diagnosis. *Lung cancer* 2010;**68**: 216-21.

7. Lopes Pegna A, Picozzi G, Mascalchi M, Maria Carozzi F, Carrozzi L, Comin C, Spinelli C, Falaschi F, Grazzini M, Innocenti F, Ronchi C, Paci E, et al. Design, recruitment and baseline results of the ITALUNG trial for lung cancer screening with low-dose CT. *Lung cancer* 2009;**64**: 34-40.

8. Lopes Pegna A, Picozzi G, Falaschi F, Carrozzi L, Falchini M, Carozzi FM, Pistelli F, Comin C, Deliperi A, Grazzini M, Innocenti F, Maddau C, et al. Four-year results of low-dose CT screening and nodule management in the ITALUNG trial. *J Thorac Oncol* 2013;**8**: 866-75.

9. Horeweg N, van Rosmalen J, Heuvelmans MA, van der Aalst CM, Vliegenthart R,Scholten ET, ten Haaf K, Nackaerts K, Lammers JW, Weenink C, Groen HJ, van OoijenP, de Jong PA, de Bock GH, Mali W, de Koning HJ, Oudkerk M. Lung cancerprobability in patients with CT-detected pulmonary nodules: a prespecified
analysis of data from the NELSON trial of low-dose CT screening. Lancet Oncol.2014 Nov;15(12):1332-41. doi: 10.1016/S1470-2045(14)70389-4.

10. Paci E, Puliti D, Lopes Pegna A, Carrozzi L, Picozzi G, Falaschi F ,Pistelli F, Aquilini F, Ocello C, Zappa M, Carozzi Fm, Mascalchi M and the Italung Working Group. Mortality, survival and incidence Rates in the Italung Randomised Lung Cancer Screening Trial. Thorax 2017, In Print.

11. Moyer VA, Force USPST. Screening for lung cancer: U.S. Preventive Services Task Force recommendation statement. *Annals of internal medicine* 2014;**160**: 330-8.

12. Field JK, Devaraj A, Duffy SW, Baldwin DR. CT screening for lung cancer: Is the evidence strong enough? *Lung cancer* 2016;**91**: 29-35.

13. Tammemagi MC, Church TR, Hocking WG, Silvestri GA, Kvale PA, Riley TL, Commins J, Berg CD. Evaluation of the lung cancer risks at which to screen ever- and never-smokers: screening rules applied to the PLCO and NLST cohorts. *PLoS Med* 2014;**11**: e1001764.

14. Kovalchik SA, Tammemagi M, Berg CD, Caporaso NE, Riley TL, Korch M, Silvestri GA, Chaturvedi AK, Katki HA. Targeting of low-dose CT screening according to the risk of lung-cancer death. *N Engl J Med* 2013;**369**: 245-54.

15. Field JK, Duffy SW, R. BD, Whynes DK, Devaraj A, Brain KE, Eisen T, Gosney J, Green BA, Holemans JA, Kavanagh T, Kerr KM, et al. UK Lung Cancer RCT Pilot Screening Trial: baseline findings from the screening arm provide evidence for the potential implementation of lung cancer screening. *Thorax* 2016;**71**: 161-70.

16. Mascalchi M, Lopes Pegna A, Carrozzi L, Carozzi F, Falaschi F, Picozzi G, Paci E. Does UKLS strategy increase the yield of screen-detected lung cancers? A comparison with ITALUNG. *Thorax* 2016;**71**: 950-1.

17. Kim Y, Kim DH. CpG island hypermethylation as a biomarker for the early detection of lung cancer. *Methods Mol Biol* 2015;**1238**: 141-71.

18. Liloglou T, Bediaga NG, Brown BR, Field JK, Davies MP. Epigenetic biomarkers in lung cancer. *Cancer Lett* 2014;**342**: 200-12.

19. Hubers AJ, Heideman DA, Duin S, Witte BI, de Koning HJ, Groen HJ, Prinsen CF, Bolijn AS, Wouters M, van der Meer SE, Steenbergen RD, Snijders PJ, et al. DNA hypermethylation analysis in sputum of asymptomatic subjects at risk for lung cancer participating in the NELSON trial: argument for maximum screening interval of 2 years. *Journal of clinical pathology* 2016.

This article is protected by copyright. All rights reserved.

20. Sozzi G, Boeri M, Rossi M, Verri C, Suatoni P, Bravi F, Roz L, Conte D, Grassi M, Sverzellati N, Marchiano A, Negri E, et al. Clinical utility of a plasma-based miRNA signature classifier within computed tomography lung cancer screening: a correlative MILD trial study. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2014;**32**: 768-73.

21. Bianchi F, Nicassio F, Marzi M, Belloni E, Dall'olio V, Bernard L, Pelosi G, Maisonneuve P, Veronesi G, Di Fiore PP. A serum circulating miRNA diagnostic test to identify asymptomatic high-risk individuals with early stage lung cancer. *EMBO molecular medicine* 2011;**3**: 495-503.

22. Montani F, Marzi MJ, Dezi F, Dama E, Carletti RM, Bonizzi G, Bertolotti R, Bellomi M, Rampinelli C, Maisonneuve P, Spaggiari L, Veronesi G, et al. miR-Test: a blood test for lung cancer early detection. *J Natl Cancer Inst* 2015;**107**: djv063.

23. Czarnecka KH, Migdalska-Sek M, Antczak A, Pastuszak-Lewandoska D, Kordiak J, Nawrot E, Domanska D, Kaleta D, Gorski P, Brzezianska EB. Allelic imbalance in 1p, 7q, 9p, 11p, 12q and 16q regions in non-small cell lung carcinoma and its clinical association: a pilot study. *Molecular biology reports* 2013;**40**: 6671-84.

24. Xia S, Huang CC, Le M, Dittmar R, Du M, Yuan T, Guo Y, Wang Y, Wang X, Tsai S, Suster S, Mackinnon AC, et al. Genomic variations in plasma cell free DNA differentiate early stage lung cancers from normal controls. *Lung cancer* 2015;**90**: 78-84.

25. Liloglou T, Maloney P, Xinarianos G, Fear S, Field JK. Sensitivity and limitations of high throughput fluorescent microsatellite analysis for the detection of allelic imbalance: application in lung tumors. *International journal of oncology* 2000;**16**: 5-14.

26. Field JK, Liloglou T, Xinarianos G, Prime W, Fielding P, Walshaw MJ, Turnbull L. Genetic alterations in bronchial lavage as a potential marker for individuals with a high risk of developing lung cancer. *Cancer Res* 1999;**59**: 2690-5.

27. Mascalchi M, Picozzi G, Falchini M, Vella A, Diciotti S, Carrozzi L, Pegna AL, Falaschi F. Initial LDCT appearance of incident lung cancers in the ITALUNG trial. *European journal of radiology* 2014;**83**: 2080-6.

28. Aberle DR, DeMello S, Berg CD, Black WC, Brewer B, Church TR, Clingan KL, Duan F, Fagerstrom RM, Gareen IF, Gatsonis CA, Gierada DS, Jain A, Jones GC, Mahon I, Marcus PM, Rathmell JM, Sicks J; National Lung Screening Trial Research Team. Results of the two incidence screenings in the National Lung Screening Trial. NEngl J Med. 2013 Sep 5;369(10):920-31. doi: 10.1056/NEJMoa1208962.

29. Shaw P, Pepe M, TA. A, Etzioni R. Methods for Assessing Improvement in Specificity when a Biomarker is Combined with a Standard Screening Test. *Stat Biopharm Res* 2009;**1**: 18-25.

Accepted Article

This article is protected by copyright. All rights reserved.

Table 1. Characteristics of the subjects sampled in the ITALUNG biomarker study according to LDCT screening group stratification, mean age and 95% Confidence Interval (C.I.), gender, active smoking, and pack years.

	n	Age (mean and 95% C.I.)	Gender (%)		Smokers at baseline (%)	Pack-years (mean and 95% C.I.)
			Male (n = 311)	Female (n = 206)		
S1 group [a]	235	60.9 (60.3-61.4)	57.9	42.1	65.5	41.3 (38.9-43.8)
S2 group [b]	118	61.1 (60.3-61.9)	57.6	42.4	67.8	43.7 (40.2-47.1)
S3 group [c]	128	60.9 (60.1-61.7)	61.7	38.3	71.9	50.5 (41.3-59.8)
S4 group [d]	18	62.9 (61.2-64.7)	68.4	31.6	78.9	54.6 (35.6-73.6)
S5 group [e]	18	63.6 (61.3-65.9)	88.2	11.8	64.7	54.7 (44.0-65.4)
Total	517	61.1 (60.7- 61.5)	60.2	39.8	68.1	43.7 (41.7-45.7)

[a] Always negative for Non Calcific Nodule.
[b] Baseline negative, positive for Non Calcific Nodule at an annual repeat test.
[c] Baseline positive for Non Calcific Nodule.
[d] Lung cancer at baseline test.
[e] Lung cancer at annual repeat test.

Accepted Article

This article is protected by copyright. All rights reserved.

Table 2. Positivity of the sputum and/or plasma genomic instability and plasma DNA, by LDCT screening group stratification.

	S1 group [a] (n = 235)	S2 group [b] (n = 118)	S3 group [c] (n = 128)	S4 group [d] (n = 18)	S5 group [e] (n = 18)
	n (%)				
Genomic instability [f]	81 (34.5%)	45 (38.1%)	42 (32.8%)	16 (88.9%)	12 (66.7%)
Plasma DNA [g] (cut-off 5 ng DNA/ml plasma)	35 (18.0%)	12 (11.9%)	17 (14.0%)	12 (66.7%)	5 (29.4%)

[a] Always negative for Non Calcific Nodule.
[b] Baseline negative, positive for Non Calcific Nodule at an annual repeat test.
[c] Baseline positive for Non Calcific Nodule.
[d] Lung cancer at baseline test.
[e] Lung cancer at annual repeat test.
[f] Positivity was defined when a subject was positive for microsatellite alteration in sputum and/or plasma.
[g] Some figures do not add up to the total because of 66 missing values.

Accepted Article

This article is protected by copyright. All rights reserved.

Table 3. The biomarker ITALUNG study sample by LDCT and IBP screening result at baseline. Number of subjects (row %).

LDCT		IBP		
		Positive	**Negative**	
	Subgroup	**n (row%)**	**n**	**n (col%)**
Positive	**S3** [a]	48 (37%)	80	128
	S4 [b]	17 (94%)	1	18
	Sum	65 (44%)	81	146 (28%)
Negative	**S1** [c]	97 (41%)	138	235
	S2 [d]	47 (40%)	71	118
	S5 [e]	12 (66%)	6	18
	Sum	156 (42%)	215	371 (72%)
		221	296	517 (100%)

[a] Baseline positive for Non Calcific Nodule.
[b] Lung cancer at baseline test.
[c] Always negative for Non Calcific Nodule.
[d] Baseline negative, positive for Non Calcific Nodule at an annual repeat test.
[e] Lung cancer at annual repeat test.

Accepted Article

This article is protected by copyright. All rights reserved.

Accepted Article

Table 4. Simulation of the results of IBP screening tests in the ITALUNG screened arm by LDCT at baseline.

LDCT	IBP		
	Positive	Negative	
	n (col%) (row%)	n (col%) (row%)	n (col%) (row%)
Positive	170 (31%) (40%)	250 (29%) (60%)	420 (30%) (100%)
Negative	378 (69%) (38%)	608 (71%) (62%)	986 (70%) (100%)
Total	548 (100%) (39%)	858 (100%) (61%)	1406 (100%) (100%)

This article is protected by copyright. All rights reserved.

Accepted Article

Table 5. All the LDCT lung cancer screen detected and interval lung cancers in the ITALUNG screened arm, by IBP screening test result.

LDCT at baseline	IBP at baseline		Total
	Positive	Negative	
	n	n	n
Positive	17 [a]	1 [a]	18 [a]
Negative	14 [b]	9 [c]	23 [d]
	31	10	41 [e]

[a] Lung cancers screen detected at baseline.
[b] One interval cancer (2 years) at baseline, 12 lung cancers screen detected at repeated round, and one interval lung cancer at repeated round.
[c] One interval cancer (2 years) at baseline, 6 lung cancers screen detected at repeated round, and 2 interval lung cancers at repeated round.
[d] Two interval cancer (2 years) at baseline, 18 lung cancers screen detected at repeated round, and 4 interval lung cancers at repeated round.
[e] One interval cancer (2 years) at repeat screening with missing biomarkers not included.

This article is protected by copyright. All rights reserved.

Table 6. Measures of accuracy by lung cancer screening modality at baseline in the ITALUNG screened cohort (n = 1406) and hypothetical strategies (simulation).

Screening modality at baseline	True positive	False positive	False negative	True negative	Sensitivity (%)	Specificity (%)	PPV	NPV
LDCT	18 [a]	402	2 [b]	984	90.0%	71.0%	4.3%	99.8%
IBP	18 [c]	530	2 [d]	856	90.0%	61.8%	3.3%	99.8%
Hypothetical combined screening strategies								
LDCT, IBP if positive [e]	18 [a]	152	2 [b]	1236	90.0%	89.0%	10.6%	99.8%
IBP, LDCT if positive [e]	18 [c]	152	2 [d]	1236	90.0%	89.0%	10.6%	99.8%

[a] Lung cancers screen detected at baseline.
[b] Interval cancers (2 years) at baseline.
[c] 17 lung cancers screen detected at baseline, and one interval cancer (2 years) at baseline.
[d] One lung cancer screen detected at baseline, and one interval cancer (2 years) at baseline.
[e] Frequencies of true positive and false negative are equal in the two strategies, but detection modality ([sd] or [ic]) by LDCT is different.

Accepted Article

This article is protected by copyright. All rights reserved.