Recent advances of microRNA-based molecular diagnostics to reduce false-positive lung cancer imaging


Mattia Boeri1‡, Stefano Sestini2‡, Orazio Fortunato1, Carla Verri1, Paola Suatoni2, Ugo Pastorino2 and Gabriella Sozzi*1
1Tumor Genomics Unit, Department of Experimental Oncology and Molecular Medicine, Fondazione IRCCS Istituto Nazionale dei Tumori, via Venezian 1, Milan 20133, Italy
2Thoracic Surgery Unit, Department of Surgery, Fondazione IRCCS Istituto Nazionale dei Tumori, via Venezian 1, Milan 20133, Italy
*Author for correspondence: gabriella.sozzi@istitutotumori.mi.it
‡Authors contributed equally

Lung cancer is the leading cause of cancer deaths in the world. Advances in early detection crucial to enable timely curative surgery have been made in recent years. Cost–benefit profiles of lung cancer screening in smokers by low-dose computed tomography are still under evaluation. In particular, the high false-positive rates of low-dose computed tomography, together with the issue of overdiagnosis and the overall costs of screening, prompted a focus on the development of noninvasive complementary biomarkers to implement lung cancer screening. MicroRNA are a new class of blood-based biomarkers useful for early lung cancer detection and prognosis definition. Here, we discuss the seminal publications that reported circulating microRNA signatures with the greatest potential to impact clinical activity and patient care.

KEYWORDS: biomarkers • circulating microRNA • LDCT • lung cancer • screening

Background
Lung cancer is the most common cause of death from cancer worldwide, responsible for nearly one cancer death in five (1.59 million deaths, 19.4% of the total). Because of the high fatality associated with the disease, the patterns in mortality closely follow those of incidence (1.82 million cases) [1]. Despite the remarkable reduction in the prevalence of active smokers and lung cancer mortality in men achieved by the introduction of smoking regulation in developed countries, millions of former smokers remain at high risk of cancer for many years.

Improvements in clinical management of lung cancer have been modest over the last 20 years, with an overall 5-year survival rate just above 10% in Europe and 16% in the USA. Treatment failure is mainly due to the presence of metastatic disease at diagnosis, occurring in 70% of all patients, whereas in patients resected in Stage IA, the 5-year survival rate is higher than 70% [2].

Detection of lung cancer at an early stage offers the real potential to reduce mortality with new chances of cure. The outcomes of the National Lung Cancer Screening Trial (NLST) have highlighted favorable prospects for lung cancer low-dose CT screening (LDCT), but the cost benefit profile of screening is still matter of debate in the scientific community [3]. In particular, the high false-positive rates of LDCT lead to multiple screening rounds, repeated radiation exposure, the use of invasive diagnostic follow-up procedures with associated morbidity and increased time and costs. In addition, LDCT screening showed a limited impact on the more aggressive lung cancers, achieving an overall mortality reduction of only 20%. We are facing an impending wave of pulmonary nodules rising from a combination of incidental findings and the proliferation of LDCT screening programs targeting high-risk individuals for lung cancer. The management of such indeterminate pulmonary nodules is challenging given the high frequency of lung nodule detection (20–60%
In 1990, the Lung Screening Study was started, a randomized clinical trial enrolling 3318 heavy smokers, 1660 receiving LDCT and 1658 receiving CXR at baseline [4]. The first screening studies started in 1992 in Japan, where two groups showed how LDCT could better diagnose lung cancer than CXR in populations of 1369 and 3967 high-risk individuals [10,11]. In both the studies, participants underwent both LDCT and CXR: in the smaller one, 11 of the 15 cases of lung cancer detected by LDCT were CXR negative. In the larger one, the LDCT lung cancer detection rate was 0.48% for LDCT and 0.03–0.05% for CXR.

In 1992, Henschke et al. promoted in the USA the Early Lung Cancer Action Project (ELCAP) that was designed to evaluate baseline and annual repeat screening by both LDCT and CXR in 1000 lung cancer high-risk individuals [12]. At baseline, LDCT identified suspicious noncalcified nodules in 233 cases and CXR in 68 cases including 27 versus 7 lung cancers, respectively. ELCAP developed subsequently in International-ELCAP with different countries joining this experience (e.g., Europe, Israel, China, and Japan) [13]. Between 1993 and 2005, 31,567 asymptomatic participants, 40 years or older and with a history of cigarette smoking, were initially screened with a LDCT and between 1994 and 2005 a total of 27,456 subjects underwent annual screening. At the baseline LDCT examination, 405 participants were found to have lung cancer while 5 received interim diagnosis. During annual screening, lung cancer was diagnosed in 74 participants, and no cases were pointed out as interim diagnosis. Stage I frequency was 85%, and the estimated 10-year survival rate regardless of treatment was 88%. Authors concluded that 80% of lung cancer deaths were preventable through screening, and that CT screening for lung cancer was 88% effective. Considering the disappointing results for lung cancer screening using CXR and with the advent of new imaging technologies, clinical trials adopting computed tomography and in particularly low-dose x-ray spiral computed-tomography (LDCT) were launched worldwide (Table 1). LDCT is a very sensitive technology that scans the body in a spiral path, thus allowing more pictures in a shorter time. In addition, it creates more detailed three-dimensional images being able to detect also the smaller abnormalities including early lung cancer nodules (National Cancer Institute website [9]).

**LDCT screening studies**

The first screening studies started in 1990 in Japan, where two groups showed how LDCT could better diagnose lung cancer than CXR in populations of 1369 and 3967 high-risk individuals [10,11]. In both the studies, participants underwent both LDCT and CXR: in the smaller one, 11 of the 15 cases of lung cancer detected by LDCT were CXR negative. In the larger one, the LDCT lung cancer detection rate was 0.48% for LDCT and 0.03–0.05% for CXR.

In 1992, Henschke et al. promoted in the USA the Early Lung Cancer Action Project (ELCAP) that was designed to evaluate baseline and annual repeat screening by both LDCT and CXR in 1000 lung cancer high-risk individuals [12]. At baseline, LDCT identified suspicious noncalcified nodules in 233 cases and CXR in 68 cases including 27 versus 7 lung cancers, respectively. ELCAP developed subsequently in International-ELCAP with different countries joining this experience (e.g., Europe, Israel, China, and Japan) [13]. Between 1993 and 2005, 31,567 asymptomatic participants, 40 years or older and with a history of cigarette smoking, were initially screened with a LDCT and between 1994 and 2005 a total of 27,456 subjects underwent annual screening. At the baseline LDCT examination, 405 participants were found to have lung cancer while 5 received interim diagnosis. During annual screening, lung cancer was diagnosed in 74 participants, and no cases were pointed out as interim diagnosis. Stage I frequency was 85%, and the estimated 10-year survival rate regardless of treatment was 88%. Authors concluded that 80% of lung cancer deaths were preventable through screening, and that CT screening for lung cancer could be highly cost-effective.

In 2000, the Lung Screening Study was started, a randomized clinical trial enrolling 3318 heavy smokers, 1660 receiving LDCT scan, 1658 receiving a posterior-anterior view chest x-ray (CXR) for lung cancer screening: the Johns Hopkins study [4]. In all these CXR studies, no significant difference in lung cancer-specific mortality was observed. In addition, a greater evidence of early stage cancers was highlighted [8]. Considering the disappointing results for lung cancer screening using CXR and with the advent of new imaging technologies, clinical trials adopting computed tomography and in particularly low-dose x-ray spiral computed-tomography (LDCT) were launched worldwide (Table 1). LDCT is a very sensitive technology that scans the body in a spiral path, thus taking more pictures in a shorter time. In addition, it creates more detailed three-dimensional images being able to detect also the smaller abnormalities including early lung cancer nodules (National Cancer Institute website [9]).

**Lung cancer imaging modalities**

With the advent of the industrial era, the spread of smoking habits influenced a significant increase in lung cancer. A long-lasting period had to pass to observe and to validate the connection between smoking and lung cancer. Once this was recognized, different programs for early detection were started. The National Cancer Institute sponsored several randomized clinical studies to test chest x-ray (CXR) for lung cancer screening: the Johns Hopkins study [4], the Memorial Sloan-Kettering study [5], the Mayo Lung Project [6] and later on also the multicenter Prostate, Lung, Colorectal and Ovarian Cancer Screening enrolling 154,901 individuals [7]. In all these CXR studies, no significant difference in lung cancer-specific mortality was observed. In addition, a greater number of high-risk individuals) and the fact that the large majority of these lung nodules are benign (up to 96%). Thus, ruling out malignancy noninvasively by evaluation of biomarkers, along with physician experiences, would be a useful strategy to improve the diagnostic workup. In particular, blood-based biomarkers could affect significantly screening performance of LDCT through the reduction of subjects needed to be followed up and the decrease of false-positive and overdiagnosis rates of LDCT scans.

**Table 1. Low-dose computed tomography in randomized clinical trials.**

<table>
<thead>
<tr>
<th>Study</th>
<th>Randomization and participants</th>
<th>Lung cancer at baseline</th>
<th>Stage I (%)</th>
<th>Lung cancer deaths</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSS (USA-2000)</td>
<td>CT: 1660, CXR: 1658</td>
<td>30, 48</td>
<td>40</td>
<td>–</td>
</tr>
<tr>
<td>NSLT (USA-2002)</td>
<td>CT: 26,722, CXR: 26,732</td>
<td>270, 63</td>
<td>356</td>
<td>–</td>
</tr>
<tr>
<td>NELSON (Netherlands-2003)</td>
<td>CT: 7557, Control: 8265</td>
<td>70, 64</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>DLCST (Denmark-2004)</td>
<td>CT: 2052, Control: 2052</td>
<td>17, 53</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>MILD (Italy-2005)</td>
<td>CT annual: 1190, CT biennial: 1186, Control: 1723</td>
<td>11, 62</td>
<td>12</td>
<td>–</td>
</tr>
<tr>
<td>LUSI (Germany-2007)</td>
<td>CT: 2029, Control: 2023</td>
<td>22, 82</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>ITALUNG (Italy-2004)</td>
<td>CT: 1613, Control: 1593</td>
<td>20, 48</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>DANTE (Italy-2001)</td>
<td>CT: 1276, Control: 1196</td>
<td>47, 66</td>
<td>20</td>
<td>–</td>
</tr>
<tr>
<td>Depsican (France-2002)</td>
<td>CT: 330, CXR: 291</td>
<td>8, 37</td>
<td>–</td>
<td>100</td>
</tr>
</tbody>
</table>

CT: Computed tomography; CXR: Chest x-ray.
radiograph [14]. Even if the study at the beginning was drawn for a single round, it was later expanded to a second screening round. The final results at year 1 of the screening pointed out 40 lung cancer in the LDCT and 20 in the CXR arm with a cancer rate detection that in LDCT passed from 1.9% at baseline to 0.57% at year 1 and in CXR arm from 0.45% at baseline to 0.68% at year 1.

Between August 2002 and April 2004, 53,454 high-risk subjects were enrolled at 33 different US medical centers belonging to the NLST [3]. The NLST was the largest study designed and powered to answer the question whether LDCT screening trial compared with CXR screening may reduce lung cancer mortality. Subjects were randomized to undergo either LDCT (26,722 subjects) or CXR (26,732 subjects) annually for 3 years (T0, T1 and T2). Lung cancers were pointed out in 1060 volunteers in the LDCT arm (645 cases per 100,000 person-years), while in the CXR arm, 941 lung cancer cases were observed (572 cases per 100,000 person-years). Lung cancer deaths were 247 per 100,000 person-years compared with the 309 per 100,000 person-years for LDCT and CXR, respectively; resulting in 20% reduced mortality, thanks to LDCT. Since this result reached the initial declared intent on November 4, 2010, in advance with respect to the scheduled program, the accomplishment of the primary endpoint of the study was announced. A decrease of 6.7% in all cause mortality was also observed in the LDCT arm. In 2013, the NLST group published the results using in the same cohort the refined setting criteria, Prostate, Lung, Colorectal and Ovarian Cancer Screening (M2012), showing an improvement in LDCT sensitivity from 71 to 83% and positive predictive value from 3.4 to 4% [15].

In Europe, other smaller randomized studies proceeded in parallel. The biggest one is represented by the Dutch-Belgian Randomized Lung Cancer Screening trial with over 15,000 subjects [16]. Other studies were the Danish Lung Cancer Screening Trial that was launched in 2004 enrolling 4104 subjects [17]; the MILD trial (Multicentric Italian Lung Detection) that enrolled 4099 participants between 2005 and 2011 (1723 randomized to the control group, 1190 assigned to screening with annual LDCT and 1186 with biennial LDCT) [18]; the German LUSI enrolling 4052 participants randomized in LDCT versus observation [19]; the ITALUNG study with 3206 participants and the DANTE trial where 2472 subjects were randomized between LDCT and control arm [20,21]; the French Depiscan trial in which 621 participants were randomized in LDCT and CXR arms [22]. In all these studies, even though LDCT was able to detect more early stage cancers, no mortality reduction with LDCT screening was observed.

Several further single-arm observational studies were performed worldwide, but a significant benefit in lung cancer mortality was not observed [23–26]. What clearly emerged through all these LDCT screening observational studies was the significant increase of early stage (and thus resectable) lung cancer compared with both the clinical and the CXR experience. Another observation that emerged was the substantial high-rate detection of noncalcified nodules that could rise to 50% of CT examinations. Such amount of false-positive findings prompted to consider in further investigations the issues of morbidity and cost–benefit of LDCT screening.

**Major clinical issues in LDCT screening studies**

In the attempt to clarify the real efficacy of lung cancer screening, Bach et al. applied a lung cancer prediction model to three prospective single-arm LDCT screening studies: the main outcome consisted of comparison of predicted with observed number of new lung cancer cases, lung cancer resections, advanced lung cancer cases and deaths from lung cancer [27]. They pooled together 3246 participants for the analysis: 144 lung cancers were diagnosed compared with the 44.5 predicted and intervention for lung cancer were almost 10 times more than expected. Advanced lung cancer did not decrease (42 subjects observed compared with 33.4 expected) as well as lung cancer mortality (38 observed vs 38.8 expected). The authors concluded that regardless an increase in the rate of lung cancer diagnosis and treatment during screening, a reduction in advanced lung cancer or death from lung cancer were not observed. They further underlined how although an excellent survival of patients with early stage lung cancer is mandatory for a CT screening to be beneficial, nonetheless LDCT should be able to intercept the more aggressive forms of lung cancer that have an impact on mortality.

A larger overview of all European randomized lung cancer CT screening, which can account for more than 37,000 people and whose analysis of data is expected in the next years, was recently provided [28].

The results highlighted by NLST raised debate in the lung cancer community on how to contextualize the experience in National lung cancer screening and on the potential benefits and hazards (IASLC2011). Concerning overdiagnosis in NLST trial, a recent study by Patz et al. showed that 18% of patients with LDCT detected lung cancer, 22% patients with non-small cell lung cancers (NSCLC) and up to 78.9% patients having bronchioalveolar carcinoma were result of overdiagnosis leading to further examinations with remarkable percentage (1.4% in the LDCT arm and 1.6% in the x-ray arm) of comorbidities such as hemithorax, lung collapse and psychosocial consequences [29].

Ultimately, cost-effectiveness of LDCT screening is still nowadays under debate with values ranging from US$4000 per life-year gained to more than US$250,000 per quality adjusted life-year gained [30]. So far, there is not enough knowledge to prove that LDCT screening is cost–effective.

Whereas NLST demonstrated a significant reduction in lung cancer mortality using LDCT, different evaluations should be considered in drawing up guidelines for mass lung cancer screening with particular emphasis on smoking cessation that should be an integral part of the screening. In a first aspect, the simple use of predictive tools based on nodule or patients characteristics may reduce the false-positive rate [31]. In addition, new minimally invasive test using reproducible biomarkers
Biomarkers in lung cancer

Several studies have reported blood-based biomarkers for early detection of lung cancer. Most of them, however, deal with discovery and validation studies on retrospective clinical series and only few of them reached prospective screening, a phase that addresses whether screening with selected biomarkers indeed result in an overall benefit for the screened population by impacting on survival. An ultimate biomarker should impact on cancer mortality, lead to a change in treatments or outcomes and should be potentially able to distinguish aggressive from indolent disease.

So far, no diagnostic biomarker has proven useful in lung cancer clinical practice. Besides technical issues related to difficulties in protocol standardization and lack of large-scale validation in clinical trials, genetic and biological tumor heterogeneity has likely limited the successful identification of tumor-specific markers. A ground-breaking way to identify novel and more reliable biomarkers is searching for candidates by looking not only at the tumor itself but also at the interplay between the tumor and the host with the aim to identify very early changes related to the biological reactivity of the host to a developing cancer. In this respect, epigenetic markers, above all circulating microRNAs (miRNAs), could represent ideal candidates because they act as extracellular messengers of biological signals derived from the cross talk between the tumor and its surrounding microenvironment.

Circulating miRNA

MiRNAs are short noncoding RNA emerged as critical regulators of gene expression playing a key role in physiological and pathological mechanisms. miRNA genes are located in separate genetic loci or within introns and exons of genes and the deregulation of miRNA expression observed in cancer is the result of chromosomal abnormalities, mutations, as described for miR-15a and miR-16 in CLL or polymorphisms. The deregulated miRNA expression can also be due to defects in their biogenesis machinery or epigenetic changes, as altered DNA methylation. A single miRNA can regulate hundreds of downstream genes by recognizing complementary sequences in the 3'UTRs of their target miRNAs. According to their target expression in different tissues, miRNAs can act as oncogenes or tumor suppressors by regulating several biological functions, such as cellular proliferation, differentiation, migration and apoptosis and regulation of cell cycle. In lung cancer, the expression and functional role of miRNAs have been extensively studied and their contribution to lung cancer development and progression has been proven. Blood circulating miRNAs were also reported to be promising biomarkers for cancer detection and prognosis. miRNAs are released into the bloodstream by different mechanisms such as passive leakage of cellular miRNAs from broken cells or active secretion through microvesicles or protein complexes by several cell subtypes. Membrane-bound microvesicles and exosomes are detected in various body fluids such as serum, plasma, urine, bronchoalveolar fluid and saliva. Importantly, exosomal miRNAs represent a new mechanism of cell–cell communication and can be functionally transferred to

<table>
<thead>
<tr>
<th>miRNAs</th>
<th>Expression in lung cancer</th>
<th>Cellular pathways (targets gene)</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>mir-221/mir-222</td>
<td>Upregulated</td>
<td>Migration (PTEN and TIMP3)</td>
<td>[90,91]</td>
</tr>
<tr>
<td>mir-21</td>
<td>Upregulated</td>
<td>Apoptosis, proliferation and migration (TPM1, PDCD4 and PTEN)</td>
<td>[92-95]</td>
</tr>
<tr>
<td>mir-17/92a cluster</td>
<td>Upregulated</td>
<td>Proliferation and cancer development (PTEN, HIF-1a CL2L11, CDKN1A and TSP-1)</td>
<td>[96-99]</td>
</tr>
<tr>
<td>mir-155</td>
<td>Upregulated</td>
<td>Cellular apoptosis and DNA damage (APAF-1)</td>
<td>[100]</td>
</tr>
<tr>
<td>mir-34 family</td>
<td>Downregulated</td>
<td>Cell death and proliferation (BCL-2, MET, PDGFR-α/β)</td>
<td>[101-103]</td>
</tr>
<tr>
<td>mir-15a/16 cluster</td>
<td>Downregulated</td>
<td>Cell cycle regulation (cyclin D1, D2 and E1)</td>
<td>[104,105]</td>
</tr>
<tr>
<td>mir-200 family</td>
<td>Downregulated</td>
<td>Induction of EMT and metastasis (ZEB transcription factors, CDH-1, vimentin)</td>
<td>[106-111]</td>
</tr>
<tr>
<td>miRNA-29 family</td>
<td>Downregulated</td>
<td>Epigenetic regulation of gene expression (DNMT-3A and DNMT-3B)</td>
<td>[112]</td>
</tr>
<tr>
<td>Let-7 family</td>
<td>Downregulated</td>
<td>– Proliferation (KRAS, MYC, HMGA2) – miRNA maturation Dicer mediated – Cell-cycle regulation (CDC25A, CDK6 and cyclin D2)</td>
<td>[113-115,116,117]</td>
</tr>
<tr>
<td>mir-548</td>
<td>Downregulated</td>
<td>Tumor growth (CCND, ERBB2, DNMT3A, DNMT3B)</td>
<td>[118]</td>
</tr>
<tr>
<td>mir-660</td>
<td>Downregulated</td>
<td>Tumor growth, migration, invasion (MDM2)</td>
<td>[49]</td>
</tr>
<tr>
<td>mir-486-5p</td>
<td>Downregulated</td>
<td>Tumor growth, migration, invasion and cell survival (ARGHAP5, p85, Pim-1)</td>
<td>[119-121]</td>
</tr>
<tr>
<td>mir-126</td>
<td>Downregulated</td>
<td>Proliferation, drug resistance (VEGF, SLC7A5, EGFR7)</td>
<td>[122-124]</td>
</tr>
</tbody>
</table>
recipient cells. By endocytic uptake or scavenger receptors, these miRNAs mediate gene expression of targets in the recipient cells [45–49]. A functional role of miRNAs associated with exosomes in cancer progression was described. In breast cancer, exosomes derived from patients mediated silencing of target cell transcriptome inducing nontumorigenic epithelial cells to form tumors [50].

**Platforms for detection of circulating miRNA**

Because miRNAs are highly stable in plasma/serum samples [51], they can be quantified using standard assays such as quantitative PCR (qPCR), hybridization technology (HYB, i.e., microarray) or next generation sequencing (NGS) [52]. Among the several qPCR technologies available, the most adopted for miRNA analysis is the stem–loop RT followed by TaqMan PCR analysis [53]. The RT step for cDNA synthesis involves the use of stem–loop RT primers to increase RT efficiency and the specificity for the miRNA mature form avoid binding to miRNA precursors. The TaqMan PCR works with hydrolytic probes with a fluorescent reporter/quencher dyes combination. During each PCR cycle to replicate the cDNA, the Taq-polymerase extends the primer upstream of the probe, hydrolyzing the probe and thus allowing the release of the fluorescent signal by the reporter dye that is proportional to the numbers of molecules amplified [54].

The first two articles reporting on circulating miRNAs as noninvasive diagnostic markers for cancer detection [55,56] quantified miRNAs using TaqMan RT-qPCR assays. Of note, the group of Tewari was the first to introduce a preamplification step to increment assay sensitivity [56,57]. More recently, other groups reported that sensitivity and robustness of qPCR can be improved using locked nucleic acid-based platforms, which have a higher thermal stability, thus avoiding the potential biases introduced by preamplification [58–60].

Microarray technology is based on the use of multiple specific probes spotted on glass, quartz or nylon chips [61]. MiRNAs are linked to fluorescent dyes and then spotted on the chip, where they are labeled to the probes by enzymatic reaction. After a washing step to eliminate nonspecific bindings, the signal of the fluorescent dye is measured by the machine. More recently, to improve assay specificity, many platforms use locked nucleic acid probes [62], or probes with hairpin structure to avoid hybridization to miRNA precursors [63].

NGS is the newest technology available for miRNAs quantification and is able to provide the exact nucleotide sequence, thus distinguishing paralogous miRNAs differing for only one base [64]. Specific adapters are initially bound to fractioned total RNA that is reverse transcribed to generate cDNA. cDNA was then amplified by standard or emulsion PCR to generate the library that is finally sequenced [65]. The main issue for a clinical test based on NGS is the large amount of total RNA needed to start the process.

Recently, Mestdagh et al. performed a miRNA quality control (miRQC) study using these three technologies, by analyzing 196 miRNAs in tissues and serum samples and evaluating seven parameters to assess the performance of 12 commercially available platforms [66]. Platforms based on NGS and HYB technologies resulted in a better titration response and had higher reproducibility and specificity, whereas qPCR platforms were more accurate, sensitive and with a higher detection rate, especially for low input RNA samples (i.e., body fluids). However, a relevant issue of reproducibility emerged when considering the 66 miRNAs differentially expressed among samples in at least one platform. In fact, only two (3%) miRNAs were concordantly differentially expressed by all the platforms, and comparing any two platforms the average validation rate was 54.6%, thus limiting the possibility to successfully validate the results obtained using the different technologies.

Focusing on circulating miRNA biomarkers, microarrays and NGS methods are used especially in a high-throughput discovery phase. In fact, both the techniques allow to profile thousand miRNAs simultaneously and NGS has also the advantage to identify all miRNAs present in the sample, even those not yet characterized, thus allowing the discovery of new molecules [65–67]. However, for further validation analyses or hypothesis-driven miRNA selection, qPCR is generally

---

**Figure 1. Origin and release of circulating miRNAs in lung cancer patients.** The circulating miRNAs are released by different cell components of the tumor and tumor–microenvironment, mostly packed in exosomes or associated to protein complexes (i.e., Ago2).
preferred because it is cheaper, more widely used (Figure 2) and easier to be transferred into the clinical practice [51,67]. In any case, every technique has its strengths and weaknesses and should be chosen according to each experimental design and purpose. In our opinion, RT-qPCR remains the best method for the analysis of low input RNA plasma/serum samples, especially to translate the scientific discoveries into the clinical practice.

Circulating miRNA-based test for lung cancer diagnosis & prognosis

The section below summarizes the knowledge on circulating diagnostic miRNAs that could fit, by their nature, in a prevention and screening policy although not all of them are consistently validated (Table 3). Serum levels of miR-1254 and miR-574-5p were significantly increased in the early stage NSCLC samples compared with the controls with 73% of sensitivity and 71% of specificity [68]. Another study described that serum expression of miR-146b, miR-221, let-7a, miR-155, miR-17-5p, miR-27a and miR-106a was significantly reduced in NSCLC cases, while miR-29c level was significantly increased [69]. Hu et al. found a serum 4-miRNA signature (miR-486, miR-30d, miR-1 and miR-499) that was significantly associated with overall survival of NSCLC patients [70]. On the basis of their previous published work, Jang and his lab identified four plasma miRNAs, miR-21, miR-126, miR-210 and miR-486-5p, that were able to distinguish lung cancer patients from controls with 86.2% sensitivity and 96.6% specificity [71]. Furthermore, the combination of miR-155, miR-197 and miR-182 plasma expression levels

<table>
<thead>
<tr>
<th>miRNA (Technology)</th>
<th>Sample</th>
<th>Significance</th>
<th>Validation set</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>miR-1254, miR-574-5p (Microarray + qRT-PCR)</td>
<td>Serum</td>
<td>NSCLC patients (11) vs controls (11)</td>
<td>22 patients vs 31 controls</td>
<td>[68]</td>
</tr>
<tr>
<td>miR-146b, miR-221, let-7a, miR-155, miR-17-5p, miR-27a, miR-106a, miR-29c (qRT-PCR)</td>
<td>Serum</td>
<td>NSCLC patients (220) vs controls (220)</td>
<td>[69]</td>
<td></td>
</tr>
<tr>
<td>miR-486, miR-30d, miR-1, miR-499 (qRT-PCR)</td>
<td>Serum</td>
<td>Prognosis in NSCLC patients: Longer (30) vs shorter (30) survival</td>
<td>243 samples</td>
<td>[70]</td>
</tr>
<tr>
<td>34-miRNA signature (qRT-PCR)</td>
<td>Serum</td>
<td>LC patients (59) vs controls (69)</td>
<td>48 patients vs 1067 controls</td>
<td>[78,81]</td>
</tr>
<tr>
<td>miR-21, miR-126, miR-210, miR-486-5p (qRT-PCR)</td>
<td>Plasma</td>
<td>LC patients (28+58) vs controls (29)</td>
<td>[71]</td>
<td></td>
</tr>
<tr>
<td>miR-155, miR-197, miR-182 (qRT-PCR)</td>
<td>Plasma</td>
<td>LC patients (74) vs controls (68)</td>
<td>[72]</td>
<td></td>
</tr>
<tr>
<td>24-miRNA signature (qRT-PCR)</td>
<td>Plasma</td>
<td>LC patients (41) vs controls (81)</td>
<td>85 patients vs 1000 controls</td>
<td>[77,79]</td>
</tr>
<tr>
<td>miR-155, miR-21 (qRT-PCR)</td>
<td>Sputum</td>
<td>NSCLC patients (23) vs controls (17)</td>
<td>[74]</td>
<td></td>
</tr>
<tr>
<td>miR-31, miR-210 (qRT-PCR)</td>
<td>Sputum</td>
<td>LC patients (62) vs controls (68)</td>
<td>143 patients vs 148 controls</td>
<td>[83,84]</td>
</tr>
<tr>
<td>let-7f, miR-30e-3p (qRT-PCR)</td>
<td>Plasma exosome</td>
<td>Prognosis in NSCLC patients (28) vs controls (20)</td>
<td>78 patients vs 48 controls</td>
<td>[73]</td>
</tr>
<tr>
<td>miR-378a, miR-379, miR-139-5p, miR-200b-5p, miR-151a-5p, miR-30a-3p, miR-629, miR-100, miR-154-3p (qRT-PCR)</td>
<td>Plasma exosome</td>
<td>LC patients (10) vs Granulomas (10) vs Controls (10)</td>
<td>50 patients vs 30 granulomas vs 25 controls</td>
<td>[85]</td>
</tr>
</tbody>
</table>
could discriminate lung cancer patients from controls with 81.3% sensitivity and 86.8% specificity in a cohort of 74 lung cancer patients and 68 age-matched cancer-free subjects [72]. Restricting the analysis to plasma vesicle-related miRNAs, Silva et al. analyzed the expression profile of NSCLC patients and identified let-7f and miR-30e-3p as new biomarkers to discriminate patients with poor outcome [73].

Another potential noninvasive source of circulating miRNAs for early detection of lung cancer is sputum. In 2010, miR-155 and miR-21 were first demonstrated to be significantly overexpressed in sputum samples collected from NSCLC patients compared with those collected from cancer-free individuals [74]. In the sputum samples, the combination of four miRNAs (miR-21, miR-486, miR-375 and miR-200b) was able to identify lung adenocarcinoma patients from normal subjects with 80.6% sensitivity and 91.7% specificity [75].

Comparing the miRNAs identified, some were overlapping among these studies (Figure 3). In particular, a group of five miRNAs was validated in three or more studies: miR-486-5p, miR-21, miR-17-5p, miR-155 and miR-126. Interestingly, all of them have an active role in lung cancer development, enhancing a proproliferative phenotype, allowing cells to escape apoptosis, regulating cell death and survival or promoting angiogenesis (Table 2).

miRNAs deregulation to complement LDCT screening

Although LDCT is currently the standard of care for early lung cancer detection [3], it results in a general overdiagnosis of indolent nodules, thus increasing individual radiation exposure, harmful confirmatory diagnostic procedures, unnecessary surgery, overload of highly specialized medical centers and increased costs for the healthcare system [29,76]. Noninvasive circulating miRNA assays could overcome most of these problems by exploiting the synergy between the molecular and the imaging tests to reduce the number of the false-positives.

Two groups, in 2011, identified specific plasma and serum miRNA signatures comparing samples from patients and disease-free individuals collected in three independent LDCT screening trials [77,78]. Boeri et al. reported four signatures composed by reciprocal ratios among 24 miRNAs by comparing samples collected before (n = 20) and at the time (n = 19) of LDCT disease detection to those of 27 control samples belonging to the INT-IEO trial [25]. These signatures were initially validated in a subset of 88 samples collected from 22 patients and 54 controls enrolled in the MILD trial [18]. Three years later, the same group developed an miRNA signature classifier (MSC), containing the 24 miRNAs previously identified, and tested its performance in enlarged validation set composed of 85 patients and 1000 controls belonging to the MILD trial [79]. The results of this study showed that the combination of MSC and LDCT reduced LDCT false-positive rate from 19.4 to 3.7%, and that the MSC risk groups were significantly associated with survival. In addition, MSC was highly sensitive (87%) and specific (81%), and its predictive value was confirmed by time-dependency analysis.

Bianchi et al. identified a 34 miRNA signature in serum samples from 59 patients enrolled in the COSMOS trial [80] compared with 69 disease-free individuals divided in training and testing sets. Globally, the test showed an AUC of 89% in the testing set, and it was also able to rule out cancer in 79% of benign lung nodules. In addition, the 34 miRNA signature did not discriminate benign or malignant breast nodules, emphasizing the specificity of the test for lung cancer. Finally, the test did not classify preinvasive plasma samples, thus limiting the capability of the test to predict the development of the disease. Very recently, the same group refined their signature to 13 miRNAs, which was validated in an independent set of 1008 subjects enrolled in the COSMOS trial [81]. Interestingly, this signature displays overlap of five miRNAs with the MSC signature (38.5%), an encouraging finding given the well-known difficulty in validating expression signatures in different studies and given the differences in samples collection between these two studies (i.e., plasma vs serum).

A subsequent work by the Jiang group showed that plasma levels of three miRNAs (miR-21, miR-210 and miR-486-5p), measured by qPCR, were able to discriminate 250 patients with CT-detected malignant solitary pulmonary nodules from benign and disease-free smokers with 76% sensitivity and 85% specificity [82], supporting the observation that miRNA biomarkers could potentially implement CT-screening.

Later on, they also identified that mir-31 and miR-210 were differentially expressed in sputum samples of 130 lung cancer patients identified by CT scan and 141 healthy individuals. The two miRNAs were able to improve specificity of CT from 83.8 to 91.1% [83]. In a very recent article, they include mir-21 in the signature and validate them in two cohorts composed by a total of 143 malignant and 148 benign solitary pulmonary nodules, resulting in 84% positive
For personal use only.

predictive value and 81% NPV, thus promoting the clinical usefulness of the test that could help in solving the critical issue of CT overdiagnosis [84].

In 2013, a collaborative group from the University of Chieti (IT) and from the NYU School of Medicine (NY, USA) developed two tests based on miRNAs present in plasma exosomes of 135 individuals with lung cancer, granulomas and healthy smokers detected by LDCT [85]. A first four-miRNA signature (miR-378a, miR-379, miR-139-5p and miR-200b-5p) was able to distinguish patients with any nodules from healthy smokers (screening test) and a different six-miRNA signature (miR-151a-5p, miR-30a-3p, miR-200b-5p, miR-629, miR-100 and miR-154-3p) classified lung cancer from granulomas (diagnostic test) with an AUC in the validation set of 90.8 and 76%, respectively. Both these signatures should be validated in an independent screening series.

### Forthcoming challenges for further advance in population screening

The challenge for the next decade will be to bring biomarkers to the clinic in ways that are efficient and practical. In this respect, the development of molecular markers able to identify tumors in a precocious phase and to track the different aggressiveness of lung tumors, including the early metastatic cancers or even the small lesions with aggressive potential is of paramount importance. Circulating miRNA profiles that reflect the physiological or pathological conditions have the potential to overcome the limitation of single blood-based biomarkers in clinical use so far and may strongly impact early cancer diagnosis. Moreover, anticipating clinical diagnosis of 1 to 2 years could significantly change the tumor burden and improve the efficacy of systemic therapies.

In 2013 at the ‘Fondazione IRCCS Istituto Nazionale dei Tumori’ (Milan, IT), a prospective screening trial implementing LDCT with the plasma MSC test was launched, the BioMILD. In 3 years, 4000 asymptomatic heavy smokers older than 50 years will be enrolled. Volunteers undergo LDCT and blood withdrawal that is immediately processed to separate plasma for miRNA profiling using custom-made microfluidic cards. On the basis of results of the MILD trial, we classified LDCT non-calcified nodules as negative (<113 mm³), indeterminate (113–260 mm³) or positive (>260 mm³). The screening algorithm is decided according to the baseline output of LDCT and MSC (Table 4).

Because the negative predictive value of MSC and LDCT together was 99% [79], if both the tests are negative, subjects do not repeat further examination for 3 years. The diagnostic strategy is established according to the size (but also taking into account shape and location) of the suspected nodule and the risk level dictated by MSC results. Repetition of LDCT at 3 months or 1 year is considered in the presence of a positive (intermediate or high risk) MSC, even in subjects with a baseline negative LDCT. In subjects with an indeterminate or positive LDCT, high-risk MSC prompts for a closer monitoring by LDCT or the use of PET examination. If hemolyzed or poor-quality plasma samples [86] are collected (15% expected), volunteers are asked to repeat the MSC test after 3 months.

The identification of a relatively small subset of smokers with substantially higher lung cancer risk could justify the immediate implementation of immunotherapy or pharmacological preventive approaches such as smoking cessation using nicotine receptors antagonists, and future development of more effective chemoinmune preventive measures. Preventive therapies aimed at correcting the miRNA imbalance linked to a damaged lung microenvironment represent an intriguing possibility [87,88].

### Expert commentary
Lung cancer is the greatest cause of cancer-related death in Western countries, and metastasis is the most common cause of death in lung cancer patients. Most patients (~60–70%) are diagnosed after the disease has spread. The 5-year survival rate for patients with advanced disease is only around 4%. However, if detected early, the 5-year survival rate is much higher (55–75%) with many of these being cured. Only 20–30% of treated NSCLC exhibit clinically significant therapeutic responses and unfortunately no certain criteria are available to distinguish a good prognosis from a bad prognosis in patients with Stage I NSCLC.

In the past, efforts to use sputum cytology and chest radiography did not achieve lung cancer mortality reduction. For the first time, in 2011, the NLST showed that a baseline LDCT screening followed by two annual screens, in contrast to standard lung x-rays, reduced lung cancer mortality by 20% and overall mortality by 7% over a 6-year follow-up period in individuals at high risk for developing lung cancer. Nonetheless, a false-positive rate of 96.4% and an overdiagnosis global rate of

---

**Table 4. BioMILD screening trial study design.**

<table>
<thead>
<tr>
<th>Baseline results</th>
<th>Low MSC (80%)</th>
<th>Intermediate MSC (16%)</th>
<th>High MSC (4%)</th>
<th>Recall time</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDCT negative (80%)</td>
<td>LDCT + MSC (65%)</td>
<td>LDCT + MSC (13%)</td>
<td>LDCT + MSC (3%)</td>
<td>3 years</td>
</tr>
<tr>
<td>LDCT indeterminate (16%)</td>
<td>LDCT + MSC (13%)</td>
<td>LDCT + MSC (3%)</td>
<td>LDCT + MSC (&lt;1%)</td>
<td>1 year</td>
</tr>
<tr>
<td>LDCT positive (4%)</td>
<td>LDCT + MSC (3%)</td>
<td>LDCT + MSC (&lt;1%)</td>
<td>LDCT or PET + MSC (&lt;1%)</td>
<td>3 months</td>
</tr>
</tbody>
</table>

(%) expected according to MILD trial results.

---

doi: 10.1586/14737159.2015.1041377
Recent advances in lung cancer screening and early detection are providing hope for patients, and their families that have never before existed in the field of lung cancer. The potential for saving lives through lung cancer low-dose computed tomography (LDCT) screening will dramatically change the disease statistics in the coming years. On the basis of NLST trial, the US Preventive Services Task Force and many professional societies have recommended annual LDCT screening for individuals at high risk. There are a number of ongoing trials evaluating value of LDCT screening for reducing lung cancer mortality in high-risk subjects. These large studies include the Dutch Belgian Randomized Lung Cancer Screening Trial and the United Kingdom Lung Cancer Screening trial and the smaller Italian trials Dante and MILD. In addition, the Italian BioMILD study represents the first effort to include in a prospective manner the use of circulating biomarkers, a plasma miRNA assay, in the diagnostic algorithm of LDCT screened smokers’ volunteers.

The discovery of small circulating molecules such as miRNAs that show higher tissue and organ specificity compared with other biological molecules and stably circulate in blood being protected by exosomes or conjugated with protein preventing their degradation represents an unprecedented opportunity for screening purposes. In particular, the integration of host–microenvironment and tumor-related biomarkers seems the most informative, earlier and sensitive approach.

The results of the large ongoing randomized prospective trials in conjunction with computer models and biomarkers will probably represent the end of this journey and will likely improve early detection, overall survival and cure rate in lung cancer.

Financial & competing interests disclosure
G Sozzi, M Boeri and U Pastorino are co-inventors for two patent applications regarding a miRNA signature for lung cancer detection. This work was supported by grants from the Italian Association for Cancer Research (AIRC): IG research grant 10096, 1227, 11991, 10068 and the special program ‘Innovative Tools for Cancer Risk Assessment and early Diagnosis’, 5 × 1000, No.12162; Italian Ministry of Health (RF-2010). Research funding was received from Genugina, Inc. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Key issues

- Lung cancer is the first cancer killer in the world, killing more people than breast, colorectal and prostate cancers combined.
- There are a number of ongoing trials evaluating value of low-dose computed tomography screening for reducing lung cancer mortality in high-risk subjects, the real value being represented by the large randomized trials and metaanalyses of smaller studies with shared methodologies and diagnostic algorithm.
- So far, no diagnostic biomarker has proven useful in lung cancer clinical practice, and tumor heterogeneity has likely limited the successful identification of tumor-specific markers.
- Epigenetic markers, above all circulating miRNAs, represent ideal candidates because they act as extracellular messengers of biological signals derived from the cross talk between the tumor and its surrounding microenvironment.
- Blood-based biomarkers could affect significant screening performance and cost effectiveness of low-dose computed tomography through the reduction of subjects needed to be followed up and the decrease of false-positive and overdiagnosis rates of low-dose computed tomography scans.
- Circulating miRNAs represent the most advanced molecular biomarkers so far identified and if validated in large prospective screening studies will likely play a central role in lung cancer screening.
References

Papers of special note have been highlighted as:
• of interest
•• of considerable interest


miRNAs diagnostics for lung cancer

Review


54. Benes V, Castoldi M. Expression profiling of microRNA using real-time quantitative PCR, how to use it and what is available. Methods 2010;50:244-9


57. microRNAs are remarkably stable, protected and easily detectable in plasma samples. In addition, they can be used to distinguish patients with prostate cancer from healthy controls, being extremely useful biomarkers.


65. Tam S, de BR, Tsao MS, McPherson JD. Robust global microRNA expression profiling using next-generation sequencing technologies. Lab Invest 2014;94:350-8


- Analysis comparing quantitative PCR (qPCR), hybridization technology and next generation sequencing platforms showed that qPCR platforms were more accurate, sensitive and with a higher detection rate and could be recommended for low input RNA samples.


For personal use only.


