**Review Article**

**Prognostic and predictive biomarkers in early stage non-small cell lung cancer: tumor based approaches including gene signatures**

Simona Carnio, Silvia Novello, Mauro Papotti, Marco Loiacono, Giorgio Vittorio Scagliotti

University of Torino, Department of Oncology, Torino, Italy

Correspondence to: Giorgio Vittorio Scagliotti. University of Torino, Department of Oncology, S. Luigi Hospital, Regione Gonzole 10, 10043 Orbassano (Torino), Italy. Email: giorgio.scagliotti@unito.it.

**Abstract:** In early stage non-small cell lung cancer (NSCLC) large randomized trials have demonstrated that in patients with radically resected disease adjuvant chemotherapy improves 5-year survival rates. However, a customization of systemic treatment is needed to avoid treatments in patients cured by surgery alone or to justify the use of adjuvant chemotherapy in high risk patients, including those in stage IA. Recently, the possibility of identifying prognostic and predictive factors related to the genetic signatures of the tumor that could affect adjuvant and neo-adjuvant treatment choices for resectable non-small cell lung cancer (NSCLC) has been of interest. This review summarizes the current status and future opportunities for clinical application of genotyping and genomic tests in early NSCLC.

**Keywords:** Prognostic and predictive biomarkers; early stages; non-small cell lung cancer (NSCLC); gene signatures

Submitted Sep 15, 2013. Accepted for publication Oct 10, 2013.

doi: 10.3978/j.issn.2218-6751.2013.10.05

Scan to your mobile device or view this article at: [http://www.tlcr.org/article/view/1610/2334](http://www.tlcr.org/article/view/1610/2334)

**Introduction**

Surgery remains the only potentially curative treatment for early-stage non-small cell lung cancer patients (NSCLC) resulting in 5-year survival rates ranging from 77% in pathological stage IA to 23% in stage IIIA tumors (1).

Clinical trials and meta-analyses have demonstrated that in patients with early-stage NSCLC adjuvant chemotherapy improves survival (2-6) with an average benefit of 5% at 5 years and, consequently, adjuvant chemotherapy is recommended for patients with resected stage II-III NSCLC (7-9). Nevertheless, a proportion of stage I patients have poor prognosis and may benefit significantly from adjuvant chemotherapy, while some relatively good prognosis stage II patients may not share similar benefits. Therefore, new diagnostic paradigms are urgently needed to select stage I-II subjects who may take advantage from adjuvant chemotherapy and clinical trials.

The strongest clinical prognostic factors in NSCLC include stage, sex, age, and performance status (10-12), but a better individualization of treatment approaches requires a more precise understanding of the molecular features of lung cancer.

A wide array of individual molecular markers have been tested in advanced as well as early stage NSCLC for prognostic and predictive value and this review will focus on the current existing evidence to support their investigational value. Most of these molecular markers have been found to be either prognostic and/or predictive at the same time. One of the first and largest retrospective biomarker studies in 515 resected stage I NSCLC patients failed to show any significant association between survival and the expression of an extensive panel of biomarkers, including epidermal growth factor receptor (EGFR), HER2/neu, bcl-2, p53 and angiogenesis markers (13).

Lastly, the human genome project, allowed the development and clinical applications of genomic-based assays, including increasingly dense microarray platforms for global analyses of gene expression, copy number
variation, DNA methylation and microRNA and several genomic signatures have been identified and tested in early stage NSCLC for their prognostic value.

**Excision repair cross-complementation group 1**

Cisplatin inhibits replication by binding to DNA and forming platinum-DNA adducts causing strand breaks when the DNA helices unwind in preparation for replication. The nuclear excision repair (NER) family of genes is involved in repair of these DNA strand breaks (14). The excision repair cross-complementation group 1 (ERCC1) enzyme is involved in the final step of the NER pathway that recognizes and removes cisplatin-induced DNA adducts, therefore, leading to cisplatin resistance. High tumoral ERCC1 expression, therefore, predicts for cisplatin resistance.

ERCC1 activity may be assessed as protein by standard immunohistochemistry (IHC) or automated AQUA technology (15). Alternatively, it can be assessed at the mRNA level through quantitative real-time polymerase chain reaction (qRT-PCR). Currently, there is no consensus about the superiority of one approach versus the other because both techniques are rarely assessed concomitantly (16,17).

In resected NSCLC, patients with high ERCC1 expression (>50 unitless ratio) had a better survival outcome (median OS, 94.6 vs. 35.5 months; P=0.01) when compared to patients with low ERCC1 expression generating the hypothesis that an intact DNA repair mechanism may reduce the accumulation of genetic aberrations that are thought to contribute to malignant potential phenotype and therefore the risk of relapse after definitive treatment (18).

The predictive role of ERCC1 was initially assessed by RT-PCR in a series of small retrospective studies in advanced NSCLC (19,20). The median overall survival (OS) was significantly longer in patients with low ERCC1 compared to patients with high ERCC1. Subsequently, ERCC1 was investigated in a subgroup of patients enrolled in a large adjuvant chemotherapy trial (21) and, by standard immunohistochemistry, in 761 paraffin-embedded tumor samples (22). A benefit from cisplatin-based adjuvant chemotherapy was associated with the absence or low expression of ERCC1 (test for interaction, P=0.009) with a significantly prolonged disease-free survival and OS among patients with ERCC1-negative tumors (HR for death, 0.65; 95% CI, 0.50–0.86; P=0.002) as opposed to ERCC-1 positive tumors. Moreover, the prognostic value of ERCC1 was confirmed in the control group with a significantly higher 5-year OS among patients with ERCC1-positive tumors than negative ones (HR, 0.66; 95% CI, 0.49–0.90; P=0.009). The same tumors were also scored by AQUA and low ERCC1 scores were marginally prognostic (HR =0.77 for high versus low scores, P=0.10) (23) while in an additional study, ERCC1 was exclusively predictive in squamous cell carcinoma (24).

The specificity of the commonly used mouse monoclonal antibody 8F1 to ERCC1 has been extensively debated (25,26). It was also found that none of the 16 antibodies tested could distinguish among the four known ERCC1 protein isoforms (27). In the neo-adjuvant setting, mRNA ERCC1 levels in pretreatment tissue samples were correlated with the capacity to achieve an objective response following platinum-based chemotherapy (P<0.05), but not with the formation of local or distant metastases (28).

The predictive value of ERCC1 was enhanced by the concurrent evaluation of MutS homolog 2 (MSH2), a major active component of the mismatch repair system. Patients with double-negative tumors experienced a greater benefit from chemotherapy (29).

Overall these data indicate a prognostic role of ERCC1 expression while its predictive role remains to be further assessed and additional validation studies are needed.

**Breast cancer susceptibility gene 1**

The protein encoded by breast cancer susceptibility gene 1 (BRCA1) has a crucial role in DNA repair as well as in cell-cycle checkpoints and mitotic spindle assembly (30). BRCA1 sensitizes cancer cells to apoptosis induced by antimicrotubule drugs, such as taxanes and vinca alkaloids, while conferring resistance to DNA-damaging agents, including platinum agents. The potential prognostic role of a panel of nine candidate biomarkers including BRCA1 was investigated in two independent cohorts of chemotherapy naive patients with early-stage NSCLC. BRCA1 was the only independent factor affecting OS (31). In a group of patients with locally advanced NSCLC, treated with neo-adjuvant cisplatin and gemcitabine followed by surgery, those with the lowest levels of BRCA1 mRNA expression had significantly greater benefit from chemotherapy in terms of clinical and pathological downsizing and OS (32).

For its localization to sites of DNA double strand breaks, the upstream activity of the receptor-associated protein 80 (RAP-80) is required for BRCA1. In a first line study in advanced NSCLC patients with the lowest expression of both BRCA1 and RAP-80 receiving cisplatin
plus gemcitabine it was shown that RAP-80 can modulate the effect of BRCA1. In addition to a close correlation with BRCA1, RAP-80 expression was identified as an independent predictor for OS (33). In a phase II feasibility study of adjuvant chemotherapy in patients with stage II-IIIA NSCLC, those with high BRCA1 transcriptional levels received single agent docetaxel, whereas those with intermediate and low BRCA1 expression were treated with cisplatin-doublets and OS did not differ between the treatment arms (34).

A recent meta-analysis of 23 studies assessed the role of BRCA1 as a predictor of clinical outcome in platinum- and paclitaxel-based chemotherapy in NSCLC patients. In 17 platinum-based studies, low-negative BRCA1 was associated with better objective response rate [ORR] (OR =1.70, 95% CI, 1.32-2.18), longer OS and event-free survival [EFS] (HR =1.58, 95% CI, 1.27-1.97, and HR =1.60, 95% CI, 1.07-2.39 for OS and EFS, respectively). In 4 paclitaxel-based chemotherapy studies, patients with high/positive BRCA1 had better ORR (OR =0.41, 95% CI, 0.26-0.64) while OS and EFS were not evaluated because of the insufficient data available (35). Some studies reported that ERCC1 expression is closely linked to RRM1 and BRCA1 levels (32,36,37), with concordant levels in 70-80% of cases (20,38).

**Ribonucleotide reductase M1**

Ribonucleotide reductase M1 (RRM1) is the regulatory component of an essential enzyme that catalyzes the reduction of ribonucleoside diphosphates to the corresponding deoxyribonucleotides. A role for ribonucleotide reductase in DNA repair has been proposed, given the capacity of RRM1 to bind a p53-regulated paralog of RRM2 called p53R2 (39,40). Increased RRM1 predicts for decreased tumor invasiveness and metastatic potential, therefore predicting for more indolent behavior, perhaps mediated through its direct correlation with phosphatase and PTEN (phosphatase and tensin homolog) protein expression (41,42). In resected NSCLC, RRM1 protein proved prognostic, with low levels associated with a median OS of 60.2 months, compared to more than 120 months in high RRM1 tumors (43). RRM1 is a major predictor of disease response to gemcitabine, being its predominant target, as well as platinum (44). Several studies investigated the predictive value of RRM1 in patients treated with gemcitabine plus cisplatin (45,46) demonstrating that RRM1 expression was significantly and inversely correlated with disease response, though not with survival (47). A recent meta-analysis in 1,243 patients with advanced NSCLC treated with gemcitabine-based regimens concluded that low tumor RRM1 was associated with a better response rate and longer survival (48).

### Thymidylate synthase

Thymidylate synthase (TS) catalyzes the conversion of deoxyuridine monophosphate (dUMP) to (deoxy) thymidine monophosphate (TMP), which requires oxidation of tetrahydrofolate to dihydrofolate. High tumoral levels of TS have been associated with resistance to 5-FU (49-51). Retrospective and prospective data from phase III trials in advanced NSCLC have established the favorable predictive value of TS in non-squamous NSCLC treated with pemetrexed with mRNA and protein TS expression lower in non-squamous compared to squamous histology and small cell lung cancer (52,53). These findings have been recently confirmed in a large retrospective study, although there was a wide range between individual patients (54).

Similarly distinct TS expression patterns among NSCLC subtypes were observed in stage I-IIIA NSCLC (55). In two different cohorts of chemotherapy naive patients with resected early-stage NSCLC, TS was a prognostic factor with mixed results. In one study high TS mRNA (but not protein) expression, was significantly associated with adverse disease free survival (DFS) and in the other study, high TS expression as determined by AQUA but not by qRT-PCR, predicted improved OS (55,56).

**ITACA (International TAilored Chemotherapy Adjuvant)** (57) trial is a randomized phase III trial comparing adjuvant pharmacogenomic-driven chemotherapy based on ERCC-1 and TS assessment by qRT-PCR versus standard adjuvant chemotherapy in completely resected stage II-IIIA NSCLC. The molecular assessment groups patients into four different genetic profiles with patients dichotomized by high versus low expression of both ERCC-1 and TS. Within 30 to 45 days post-surgery, patients in each genetic profile are randomized to receive either a standard adjuvant chemotherapy doublet (control arm) or an experimental treatment guided by molecular determinants (tailored chemotherapy arm). The study is currently accruing patients in Italy and in Germany and more than 600 patients have been already randomized (see Figure 1).
Radically Resected II-IIIA
No prior
Chemotherapy or
Radiation Therapy
prior surgery

ERCC1 and TS
Assessment by RT-PCR

HIGH ERCC1 & HIGH TS

HIGH ERCC1 & LOW TS

LOW ERCC1 & HIGH TS

LOW ERCC1 & LOW TS

Docetaxel
Standard
Chemotherapy

Pemetrexed
Standard
Chemotherapy

Cisplatin/Gemcitabine
Standard
Chemotherapy

Cisplatin/Pemetrexed
Standard
Chemotherapy

Figure 1: Study Design of The ITACA Adjuvant Trial. Randomization is performed according to each genetic profile. Cisplatin-doublet, investigator choice is the Control Arm of the study. Chemotherapy is either Vinorelbine 25-30 mg/m² IV over 10 minutes, days 1 and 8 & Cisplatin 75 mg/m² IV over 60 minutes, day 1, immediately following Vinorelbine or Docetaxel 75 mg/m² IV over 60 minutes, day 1 & Cisplatin 75 mg/m² IV over 60 minutes, day 1, immediately following Docetaxel or Gemcitabine 1,200 mg/m² IV over 30 minutes, days 1 and 8 & Cisplatin 75 mg/m² IV over 60 minutes day 1, immediately following gemcitabine. For the purpose of the final statistical analysis at the end of the study all controls will be grouped together in one single group (control group, blue boxes and line) and all tailored chemotherapies will be assembled together (experimental group, orange boxes and line).

Cyclin-dependent kinase inhibitor 1B

The cyclin-dependent kinase inhibitor 1B, p27<sup>Kipl</sup>, is a tumor-suppressor protein that induces cell-cycle arrest in phase G1. Despite its anti-proliferative properties, p27<sup>Kipl</sup> up-regulation leads to de novo resistance to platinum agents by allowing cancer cells to repair DNA damage and avoid apoptosis. A survival benefit from cisplatin-based chemotherapy was only demonstrated in patients with p27<sup>Kipl</sup> negative tumors (58,59). Cyclin D2 has been associated with poor recurrence-free survival in patients in stage III NSCLC treated with surgery with or without adjuvant chemotherapy (60).

β-tubulin

β-tubulin (βTubIII) is an essential element of microtubules, which serves as a cellular structural component involved in vital processes, including mitosis. Class IIIβ-tubulin (βTUBIII) corresponds to an isotype with enhancing impact on microtubule dynamics, contributing to intrinsic cancer cell resistance to antimitotic agents. Several studies have shown that βTubIII expression may predict response and outcome in patients with advanced NSCLC receiving tubulin binding agents (61,62). In patients with early stage NSCLC treated with an adjuvant vinorelbine-based regimen (2), high βTUBIII expression was shown to be an independent adverse predictor of recurrence-free survival (63). The prognostic value was confirmed retrospectively in patients enrolled in another adjuvant study (64) and more recently in a neo-adjuvant study (65). In patients with β-tubulin positive immunostaining, median PFS was 30.6 versus 60.1 months (HR =1.46; 95% CI, 1.08-1.99) for those negative.
**RAS oncogene and p53**

KRAS mutation was associated in some studies with a poor prognosis in early NSCLC patients (66) but not in others. In the phase III NCIC JBR.10 adjuvant trial, where patients with stage IB-II disease were prospectively stratified by the presence of mutation in any of the RAS genes, the effect of RAS mutation status on the treatment outcome and prognosis was not significant. Nevertheless, the lack of benefit from the cisplatin/vinorelbine combination in patients with RAS mutations, in contrast to the total study population, may suggest a possible negative predictive role (67). Similarly, another adjuvant study in stage IB patients showed that, the presence KRAS mutations trended for an inferior survival of patients with tumors larger than 4 cm (68). Consistent data were reported in an Italian adjuvant study where in a subgroup of 227 patients the presence of K-RAS mutation, but not p53 and Ki67, in the univariate, but not in the multivariate analysis, was associated with shorter survival (69).

p53 nuclear immunoreactivity is considered a surrogate marker of TP53 gene mutations. However, the sensitivity and positive predictive value of p53 IHC expression for TP53 mutation status are estimated to be only 75% and 65%, respectively (70). A retrospective analysis of the phase III NCIC-JBR.10 adjuvant trial showed p53 IHC overexpression to be an independent unfavorable prognostic factor among patients in the observation arm. In contrast to p53 expression, TP53 mutation status was neither prognostic for survival, nor predictive for efficacy of adjuvant chemotherapy (71).

**Gene expression signatures**

Microarray technologies allow exploration of the prognostic significance of thousands of markers using high-throughput and computational approaches. To date, in lung cancer more than 30 studies have been reported (72) a large number showing that gene expression signature may stratify early stage NSCLC patients with different prognosis or survival outcome.

Although most of these signatures have been validated in one or more independent patient cohorts, microarray dataset overlaps between the genes sets have consistently been minimal. Thus there is a strong possibility that sample collection methods, processing protocols, single-institution subject cohorts, small sample sizes, and peculiarities of the different microarray platforms are contributing significantly to the results. To address these issues, a multi-institutional collaborative study was conducted to generate gene expression profiles from a large number of samples with *a priori* determined clinical features, useful to evaluate proposed prognostic models for potential clinical implementation. A large series of lung adenocarcinomas were tested for whether microarray measurements of gene expression either alone or combined with basic clinical covariates (stage, age, sex) can be used to predict overall survival in lung cancer subjects. Risk scores were produced substantially correlated with actual subject outcome, especially when clinical and molecular information are combined to build prognostic models for early stage lung cancer (73).

A malignancy-risk gene signature composed of several genes associated with proliferative activity has been successfully applied to predict breast cancer risk (74,75) and also tested for prognostic and predictive value in an early-stage NSCLC patients. The malignancy-risk gene signature was tested using a large NSCLC microarray dataset from the Director's Challenge Consortium (n=442) and two independent NSCLC microarray datasets (n=117 and 133 datasets, respectively). The malignancy-risk gene signature was significantly associated with OS (P<0.001) in two independent datasets and also with adjuvant chemotherapy (P=0.02) (76). Xu *et al.* developed an empirical model which is not based on the knowledge of patients’ survival time for determining the lung cancer biomarker signature. It has been hypothesized that instead of an individual gene, two functionally imbalanced groups of genes (Yin and Yang) determine the fate of the tumor cells, which ultimately determines patient's survival time. The Yin and Yang genes were selected by comparing expression data from normal lung and lung cancer tissue samples using both unsupervised clustering and pathways analyses. The model was tested in four independent lung cancer datasets and significantly stratified patients into high- and low-risk survival groups and predicted chemotherapy outcomes for stages II and III (77).

A 14-gene expression assay that uses quantitative PCR, runs on formalin-fixed paraffin-embedded tissue samples, and differentiates patients with heterogeneous statistical prognoses was developed in a cohort of 361 resected patients with non-squamous NSCLC and validated in 2 different cohorts of 433 patients with resected stage I non-squamous NSCLC and 1,066 patients with stage I-III non-squamous NSCLC resected in several leading Chinese cancer centres. The signature significantly segregated patients in low-, intermediate- and high-risk patients with relevant differences in 5-year survival rate. Multivariate
analysis in both cohorts indicated that no standard clinical risk factors could account for, or provide, the prognostic information derived from tumour gene expression. This quantitative-PCR-based assay reliably identified patients with early-stage non-squamous NSCLC at high risk for mortality after surgical resection (78).

**Other biomarkers in early NSCLC**

Insulin receptor (IR) and Insulin-like growth factor receptor (IGF1R) are implicated in the development and progression of NSCLC, either by interacting with the EGFR pathway or independently. In patients with resected NSCLC, IGF1R amplification determined by FISH was an independent favorable prognostic factor, unlike IGF1R protein (79). It has been observed that in early stage NSCLC, IGF1R/EGFR FISH+ and IGF1R/EGFR IHC+ were associated with shorter disease-free survival (P=0.05 and P=0.05, respectively). Patients with concomitant IGF1R/EGFR FISH+/IHC+ had a worse DFS and OS (P=0.005 and P=0.01, respectively) (80).

Hepatocyte growth factor receptor (c-MET) is a proto-oncogene associated with tumor invasive growth. In patients with resected stage I-III carcinomas, not treated with adjuvant chemotherapy, a high MET gene copy number was an independent adverse prognostic factor mainly in squamous histotype (81). Similarly, in patients with early-stage NSCLC HER2 expression was associated with poorer prognosis especially in stages IB and IIA diseases (82).

Higher expression of CXCR7 is associated with metastatic progression and poor DFS in patients with stage I NSCLC (83). In a retrospective analysis of completely resected stage I tumors, patients with CXCR4-positive tumors had a significantly longer survival than patients with CXCR4-negative tumors (P=0.039). Interestingly, the 5-year metastasis rates were 23.5% and 34.1% in patients with CXCR4-positive and CXCR4-negative expression, respectively (P=0.2) (84).

Recently, a set of genes with altered methylation status were identified in stage I NSCLCs, some of which associated with survival. Such newly identified potential candidates for a NSCLC molecular screening need further analysis in order to determine their clinical utility (85,86).

Studies in early stage NSCLC have reported an association between vascular endothelial growth factor (VEGF) over-expression and progression or poor survival (87,88), but overall the prognostic role of VEGF expression in NSCLC remains undetermined.

**Conclusions**

Currently, tumor stage remains the strongest predictor of survival in NSCLC. Early-stage patients are treated primarily by surgical resection. However, 30% to 55% of these patients develop recurrence and die of their disease. The current staging system is inadequate for predicting the outcome of treatment and the prognosis in individual patients. For this reason there is an urgent need to search for new individual biomarkers and gene signatures in the tumor tissue. Many studies have investigated several molecular alterations in early lung cancer and their predictive and prognostic implications are still a matter of clinical research and even for the most promising are not yet ready for prime time application. Currently, randomized clinical trials are exploring the real value of these new diagnostic tools.

**Acknowledgements**

*Disclosure: The authors declare no conflict of interest.*

**References**


