Introduction

Approximately 25-30% of patients with non-small cell lung cancer (NSCLC) present with early stage disease and undergo surgery with curative intent. Despite complete tumor resection, many of these patients will develop systemic relapses with or without local relapses and will eventually die. A meta-analysis of early trials indicated a trend towards improved survival for adjuvant platinum-based chemotherapy and led to a re-evaluation of adjuvant treatment in clinical trials in large patient populations. Several of these trials demonstrated an improved survival with the use of adjuvant chemotherapy. The survival benefit was then further confirmed in a meta-analysis that included all five cisplatin-based trials (Table 1) (1-7). Early NSCLC comprises a heterogeneous group of diseases, with diverse innate aggressiveness and degree of response to cytotoxic agents. For instance, some patient subsets with stage II or even stage IIIA have excellent prognosis...
and could be spared the toxicity of unnecessary therapy. Others, such as the elderly or less fit, as well as those with stage I disease, remain undertreated despite potential benefit from adjuvant chemotherapy. This underscores the need for a customized approach to tailor adjuvant chemotherapy according to patient characteristics and tumor characteristics. Personalizing therapy based on an individual patient’s molecular profile is a potentially promising approach to optimize efficacy with available agents. Prognostic biomarkers indicate the natural course of disease, irrespective of treatment, while those defined as predictive can foresee differential therapeutic outcomes. However, some biomarkers combine both of these functions, such as ERCC1 and RRM1. Identification and application of the appropriate biomarkers would enable selection of only high-risk patients to receive the most effective treatment. In this review we describe potential predictive and prognostic markers and their current role, benefit, and possible future use in the management of patients with early stage NSCLC.

**Clinicopathological prognostic and predictive factors in early stage NSCLC**

It has long been recognized that differences in clinical factors such as stage, sex, and tumor factors such as cellular differentiation, vascularity, and vascular invasion, are prognostic of outcome and important in determining adjuvant therapy decisions for early stage NSCLC. To date, pathological stage, as defined by tumor size and nodal status, is the only prospectively validated clinicopathological biomarker with both prognostic and predictive value. According to guidelines from the European Society of Medical Oncology (ESMO), American Society of Clinical Oncology (ASCO), National Comprehensive Cancer Network (NCCN) and the American College of Chest Physicians (ACCP), platinum-based adjuvant chemotherapy is considered standard treatment for resected stage II-IIIA disease with an estimated survival benefit of 4-5% at 5 years (8-11). Although pre hoc subgroup analyses have shown no benefit for patients with pathological stage IB disease, interpretation of these results should be cautious, since the test for interaction between treatment effect and stage was not significant in any of the phase III trials with platinum-based regimens (2-4,12). Only one of the available meta-analyses did demonstrate a significant differential treatment effect, largely in the stage IA subgroup, suggesting that patients with stages II and IIIA have greatest benefit (2). The reported 11% 5 years survival gain of adjuvant tegafur-uracil in stage I NSCLC cannot be directly extended to Western countries where tegafur-uracil has not yet been reliably tested (13). At the same time, the assumption of therapeutic benefit for stage IB disease with tumor size larger than 4 cm is based on an unplanned subgroup analysis (5). Nonetheless, worse prognosis observed with increasing T size has been recognized in the 7th TNM edition. T2 was divided into T2a (3-5 cm) and T2b (5-7 cm), with 5 year overall survival of 58% and 49%, respectively (P<0.0001); T2bN0 was upstaged to stage IIA. Correlation with the

<table>
<thead>
<tr>
<th>Table 1 Adjuvant chemotherapy of completely resected NSCLC</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
</tr>
<tr>
<td>----</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>ALPI-EORTC (1)</td>
</tr>
<tr>
<td>IALT (2)</td>
</tr>
<tr>
<td>NCIC CTG-JBR.10 (3)</td>
</tr>
<tr>
<td>ANITA (4)</td>
</tr>
<tr>
<td>CALGB 9633 (5)</td>
</tr>
<tr>
<td>BLT (6)</td>
</tr>
</tbody>
</table>

ALPI, adjuvant lung project Italy; NCIC CTG-JBR.10, National Cancer Institute of Canada Clinical Trials Group-JBR.10 trial; ANITA, Adjuvant Navelbine International Trialist Association; CALGB, Cancer and Leukemia Group B; BLT, Big Lung Trial; IALT, International Adjuvant Lung Cancer Trial; LACE meta-analysis, Lung Adjuvant Cisplatin Evaluation meta-analysis; CT, chemotherapy; MVP, mitomycin C, vindesine and cisplatin; Vinca, etoposide or vinorelbine or vinblastine or vindesine; HR, hazard ratio.
new staging system failed to validate the 5 cm cut-off in the 9-year update of CALGB 9633, showing a trend towards a significant benefit for adjuvant treatment for patients with tumors >7 cm (HR=0.53; P=0.051), although this interaction should be investigated further (14-16).

With regard to the importance of accurate staging, a recent retrospective analysis of lymph node dissection in more than 20,000 patients with pathological stage I NSCLC implied that the number of recovered lymph nodes might be predictive for survival outcomes, although this could well be attributed to a direct therapeutic effect (17). Furthermore, a large retrospective analysis from the SEER database showed that the increasing number of resected positive nodes and a higher ratio between metastatic and overall resected nodes has an independent negative prognostic impact for overall survival in N1 patients (18,19). Recent studies investigated high tumor grade, vascular invasion and visceral pleural infiltration as poor prognostic determinants, based on mostly retrospective cohort studies, and are commonly recommended as adjunct selection criteria for patients who are borderline candidates for adjuvant chemotherapy (20,21).

Other clinicopathological features prospectively shown to be independent, unfavorable prognostic factors in early NSCLC include older age, male sex and non-squamous-cell histology. Only performance status was likely to predict therapeutic effect (3,4,7). Likewise, certain histological subtypes, such as large-cell neuroendocrine carcinoma and pure bronchioloalveolar carcinoma which confer, respectively, worse or better outcome, could also guide treatment strategy (22,23). Interestingly, a high correlation between gene expression profile signatures and tumor histological phenotype has been shown for early NSCLC (24). It seems that even a thoroughly validated molecular signature does not outperform combined conventional clinical and pathologic variables in predicting survival of NSCLC patients (24). Therefore, incorporating the subtype and grade into conventional clinical models could provide predictive accuracy similar to that of well validated gene panels (24).

Molecular prognostic and predictive markers in early stage NSCLC

Gene expression profiling signatures

Gene expression profiles may facilitate treatment decisions in lung cancer, similar to their use to predict chemotherapy benefit in early stage breast cancer. Several groups have developed prognostic signatures based on mRNA, microRNA or proteomic profiles in order to better define patients with good prognosis who could potentially be spared adjuvant treatments, and those with poor prognosis who may benefit from successful adjuvant therapies. However, most prognostic signatures have some limitations that should be taken into account when analyzing their potential clinical utility. For instance, survival of NSCLC patients depends to a large extent on co-morbidity factors and the impact of this cannot be accounted for by prognostic tests based on tumor molecular profiling. At the same time, methodological or statistical data analyses have often been insufficient in the original studies proposing the signature for clinical use. Consider the recent example of the phase III Cancer and Leukemia Group B (CALGB) 30506 trial that was originally designed to validate the potential utility of a lung cancer metagene model in selecting patients with pT1 -T2N0 tumors for adjuvant chemotherapy. The study was recently amended as the original authors failed to replicate their own results. This highlights the need for very careful large-scale validation of prognostic signatures before they can be prospectively tested in clinical studies with adjuvant therapies for lung cancer (25).

Subramanian and Simon have recently published an elegant review of 16 published studies involving the analysis of gene expression data for developing prognostic signatures in NSCLC, in which they report serious methodological flaws in design and analysis, including inappropriate patient dataset selection, lack of independent validation, biased reporting of re-substitution statistics, incomplete protocol specification and use of statistical methods (26). Indeed, only a few gene signature studies have yielded data specifically referring to NSCLC stages IA, IB, or II that warrant further prospective validation (27-33). Among them it is worth mentioning the prognostic 15-gene signature for early NSCLC that was recently reported as the first deriving from prospectively collected tumor samples from patients enrolled in a phase III adjuvant trial. Gene expression profiling was conducted on mRNA from 133 frozen tumor samples from the National Cancer Institute of Canada Clinical Trials Group (NCIC CTG)-JBR.10 trial (28). The prognostic value of this gene signature was tested in four independent published microarray data sets and by quantitative reverse-transcriptase polymerase chain reaction (RT-qPCR). Among these genes were
nuclear proteins or transcription regulators such as mouse double minute 2 homolog (MDM2), zinc finger protein 236 (ZNF236), fos-related antigen 2 (FOSL2), hexamethylene bis-acetamide (HMBA)-inducible protein 1 (HEXIM1), myelin transcription factor 1-like (MYT1L) and inhibitor of kappa light polypeptide gene enhancer in B-cells (IKBKAP). The second subset of genes included protein melan-A (MLANA), ATPase subunit beta-1 (ATP1B1), L1 cell adhesion molecule (L1CAM), and stathmin-2 (STMN2), which encode for transmembrane-or membrane-associated proteins, potentially involved in signaling pathways and, finally, sodium/potassium-transporting ATPase subunit beta-1 (ATP1B1) and uridine monophosphate synthetase (UMPS) which are involved in purine and pyrimidine metabolism, respectively, suggesting dependency of NSCLC on these pathways (28).

This 15-gene expression profile was unique in that it could also predict response to systemic chemotherapy, whereas most other gene profiles have served only as prognostic markers following surgery. The signature was shown to interact significantly with the effect of cisplatin plus vinorelbine chemotherapy, with high-risk patients benefiting the most, although its potential predictive role requires independent validation. Also clinically relevant to the adjuvant strategy, this signature was able to assign, separately, stage IB and II patients to high- and low-risk subgroups with significantly different overall survival (28). When the predictive value of previously published prognostic signatures, applied to treated and untreated patients in the JBR.10 data set was evaluated, only the six-gene signature identified by Boutros was proved to be both significantly prognostic and predictive (27). The six-gene model comprised: syntaxin 1A (STX1A), hypoxia inducible factor 1A (HIF1A), chaperonin containing TCP1 subunit 3 (CCT3), MHC class II DP beta 1 (HLA-DPB1), v-maf musculoaponeurotic fibrosarcoma onco-gene homolog K (MAFK), and ring finger protein 5 (RNF5) (27).

Finally, a 14-gene assay that uses RT-qPCR analysis of formalin-fixed, paraffin-embedded tissues was developed with a cohort of 361 patients with non-squamous NSCLC resected at the University of California, San Francisco, by Kratz, He, et al. from Michael Mann and David Jablons’ group (34). The investigators developed a 14-gene signature panel, consisting of 11 cancer-related genes: BCL2-associated athanogene (BAG1), breast cancer susceptibility gene 1 (BRCA1), cell division control protein 6 homolog (CDC6), cyclin-dependent kinase 2 associated protein 1 (CDK2AP1), receptor tyrosine-protein kinase erbB-3 (ERBB3), galactoside 3(4)-L-fucosyltransferas (FUT3), interleukin 11 (IL11), lymphocyte-specific protein tyrosine kinase (LCK), Rho family GTPase 3 (RND3), SH3 domain-binding glutamic acid-rich protein (SH3BGR), and wingless-type MMTV integration site family, member 3A (WNT3A) together with 3 reference genes, esterase D (ESD), TATA box binding protein (TBP) and Yes-associated protein 1 (YAP1). They validated the candidate gene signatures in 2 different populations: a community-based series of 433 resections for stage I non-squamous NSCLC from Northern California, and a cohort of 1,006 resections for stage IA-IIIB non-squamous NSCLC from the China Clinical Trials Consortium (34). The combination of gene signatures proved to be independently prognostic, irrespective of TNM stage grouping (34), in stage I, II and III patients. The prognostic value was significantly greater than certain clinical risk stratification criteria proposed by the US National Comprehensive Cancer Network for stage I resections (34). Furthermore, it was similarly effective in the Northern California and Chinese validation populations (34). However, there are some limitations to this study that should be pointed out. For instance, patients with squamous cell histology were excluded and there was poor overall quality of pathologic nodal staging, bearing in mind that 18% of resections for NSCLC in United States have no lymph nodes examined. As with all studies so far, this is another retrospective series, albeit the largest and most rigorously validated one performed to date.

Individual prognostic and predictive biomarkers in early-stage NSCLC

There are several candidate markers for sensitivity or resistance to chemotherapy identified in retrospective analyses of tumor biopsies from phase III clinical trials testing the value of adjuvant chemotherapy (Tables 2, 3).

Excision repair cross complementation group 1 (ERCC1)

ERCCI is a rate limiting enzyme in the nucleotide excision repair (NER) and interstrand cross-link repair pathways, which recognizes and repairs platinum induced adducts. Cancer cells overexpressing ERCC1 are more likely to have de novo resistance to cisplatin and a growing list of reports links cisplatin, carboplatin, and oxaliplatin resistance to ERCC1 mRNA levels in tumors. This relationship has been suggested for patients with gastric, bladder, ovarian, colorectal, and lung cancer. It was shown that ERCC1 levels evaluated by immunohistochemistry (IHC) are also
### Table 2 Prognostic biomarkers in early stage non-small cell lung cancer

<table>
<thead>
<tr>
<th>Study</th>
<th>Stage of disease</th>
<th>Number of patients</th>
<th>Study design</th>
<th>Biomarker (assay)</th>
<th>Biomarker status</th>
<th>HR for overall survival (P value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Olaussen et al., 2006 (35)</td>
<td>I-III</td>
<td>1867/761</td>
<td>Retrospective analysis within IALT study</td>
<td>ERCC1 (IHC)</td>
<td>Positive expression (H-score &gt; median value)</td>
<td>0.66 (0.009)</td>
</tr>
<tr>
<td>Kamal et al., 2010 (36)</td>
<td>I-III</td>
<td>1867/673</td>
<td>Retrospective analysis within IALT study</td>
<td>MSH2 (IHC)</td>
<td>High expression (H-score=3)</td>
<td>0.66 (0.01)</td>
</tr>
<tr>
<td>Tsao et al., 2007 (37)</td>
<td>IB-II</td>
<td>482/253</td>
<td>Retrospective analysis within NCIC CTGJBR.10</td>
<td>p53 (IHC)</td>
<td>Positive expression (staining score ≥15%)</td>
<td>1.89 (0.03)</td>
</tr>
<tr>
<td>Graziano et al., 2010 (38)</td>
<td>IB</td>
<td>344/250</td>
<td>Retrospective analysis within CALBG 9633</td>
<td>p53 (IHC)</td>
<td>Positive expression</td>
<td>2.30 (0.0005)</td>
</tr>
<tr>
<td>Seve et al., 2007 (39)</td>
<td>IB-II</td>
<td>482/265</td>
<td>Retrospective analysis within NCIC CTGJBR.10</td>
<td>p53 (IHC)</td>
<td>Positive expression</td>
<td>1.72 (0.04)</td>
</tr>
<tr>
<td>Rosell et al., 2007 (40)</td>
<td>I-IIIA; IB-IIIB</td>
<td>126; 58 (validation cohort)</td>
<td>Retrospective analysis of cohort data</td>
<td>BRCA1 (RT-qPCR)</td>
<td>High expression (relative gene Expression &gt; median value)</td>
<td>1.98 (0.02); 2.4 (0.04)</td>
</tr>
</tbody>
</table>

HR, hazard ratio

### Table 3 Predictive biomarkers in early stage non-small cell lung cancer

<table>
<thead>
<tr>
<th>Study</th>
<th>Stage of disease</th>
<th>Number of patients</th>
<th>Study design</th>
<th>Biomarker (assay)</th>
<th>Biomarker status</th>
<th>HR for overall survival (P value); P value for interaction test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Olaussen et al., 2006 (35)</td>
<td>I-III</td>
<td>1867/761</td>
<td>Retrospective analysis within IALT study</td>
<td>ERCC1 (IHC)</td>
<td>Negative expression vs. positive expression (H-score &gt; median value)</td>
<td>0.65 (0.002) vs. 1.14 (0.40); 0.009</td>
</tr>
<tr>
<td>Kamal et al., 2010 (36)</td>
<td>I-III</td>
<td>1867/658</td>
<td>Retrospective analysis within IALT study</td>
<td>MSH2/ERCC1 (IHC)</td>
<td>Both low vs. both high</td>
<td>0.65 (0.01) vs. 1.32 (0.19); 0.01</td>
</tr>
<tr>
<td>Kamal et al., 2010 (36)</td>
<td>I-III</td>
<td>1867/not defined</td>
<td>Retrospective analysis within IALT study</td>
<td>MSH2/p27 (IHC)</td>
<td>Both low vs. both high</td>
<td>0.65 (0.01) vs. 1.31 (0.22); 0.01</td>
</tr>
<tr>
<td>Kamal et al., 2010 (36)</td>
<td>I-III</td>
<td>1867/673</td>
<td>Retrospective analysis within IALT study</td>
<td>MSH2 (IHC)</td>
<td>Low expression vs. high expression (H score =3)</td>
<td>0.76 (0.03) vs. 1.12 (0.48); 0.06</td>
</tr>
<tr>
<td>Tsao et al., 2007 (37)</td>
<td>IB-II</td>
<td>482/253</td>
<td>Retrospective analysis within NCIC CTGJBR.10</td>
<td>p53 (IHC)</td>
<td>Positive expression (staining score ≥15%) vs. negative</td>
<td>0.54 (0.02) vs. 1.40 (0.26); 0.02</td>
</tr>
<tr>
<td>Filipits et al., 2007b (41)</td>
<td>I-III</td>
<td>1867/778</td>
<td>Retrospective analysis within IALT study</td>
<td>p27 (IHC)</td>
<td>Negative expression vs. positive expression (H-score &gt; median value)</td>
<td>0.66 (0.006) vs. 1.09 (0.54); 0.02</td>
</tr>
<tr>
<td>Pirker et al., 2007 (42)</td>
<td>I-III</td>
<td>1867/778</td>
<td>Retrospective analysis within IALT study</td>
<td>ERCC1/p27 (IHC)</td>
<td>Both negative vs. both positive</td>
<td>0.52 (95% CI: 0.36-0.74) vs. 1.27 (95% CI: 0.87-1.84); not specified</td>
</tr>
</tbody>
</table>

HR, hazard ratio; vs., versus
predictive for the survival benefit afforded by adjuvant cisplatin-based chemotherapy in patients with totally resected stage I to IIIA NSCLC (35).

The International Adjuvant Lung Cancer Trial (IALT)-Bio translational research project aimed to study molecular biomarkers of tumors for their potential predictive values with regard to the effect of adjuvant chemotherapy on survival in IALT patients. Five groups of molecular biomarkers (19 markers in total) were studied by IHC: drug transporters, DNA repair, cell cycle regulators, signal transduction and apoptosis. Both ERCC1 and cyclin-dependent kinase inhibitor 1B (CDKN1B or p27) were found to have predictive value in patients with completely resected NSCLC undergoing adjuvant cisplatin-based chemotherapy (35). Interestingly, in patients randomly assigned to the observation arm, the subgroup with ERCC1-positive tumors had better survival compared with those with ERCC1-negative tumors (35). The paradoxical status of ERCC1, which was found to be a good prognostic marker in untreated resected NSCLC patients but a poor predictor of efficient adjuvant chemotherapy, was also confirmed in another study by Zheng et al, where the concomitant high expression of RRM1 and ERCC1 delineated a subgroup of chemonaïve patients with stage I disease with excellent survival outcomes (43). p27 is a tumor-suppressor protein that induces cell-cycle arrest in phase G1. Overexpression of p27 may confer de novo resistance to cisplatin by giving necessary time to repair cisplatin-induced DNA damage. In retrospective analysis of the IALT, patients with p27-negative tumors had longer survival after chemotherapy compared with surgery alone (41). Among six cell cycle regulators evaluated by IHC within the IALT-bio project, only p27 was identified to significantly correlate with treatment effect. Its predictive ability was independent from ERCC1 expression and, as anticipated, only patients with p27-negative tumors had survival benefit of cisplatin-based chemotherapy. Furthermore, when combining the IHC features of ERCC1 and p27, patients with tumors negative for both biomarkers seemed to benefit most from adjuvant chemotherapy (41).

In a pharmacogenomic trial with a biomarker-strategy design in advanced NSCLC by Cobo et al, ERCC1 mRNA expression was evaluated prospectively in an attempt to predict response to cisplatin-based or cisplatin-free chemotherapy regimens in stage IIIB or IV in NSCLC patients (44). Overall response rate was significantly higher in the genotypic arm, where chemotherapy regimen was tailored by ERCC1 mRNA expression (44). Patients in the control arm were not evaluated for the biomarker and received standard platinum-based combination (44). Within the customized arm, patients with low ERCC1 levels were treated with the same regimen as the control arm, whereas those with high levels received a non-platinum regimen (44). Most importantly, however, clinical relevance remained limited, given that there was no difference between the two arms in either progression-free survival or overall survival. This paradox of favorable long-term outcome despite cisplatin chemo-resistance probably indicates that, by preventing mutagenesis, DNA repair may not only prevent cancer but may retard molecular events related to progression in established tumors. Thus, high expression of ERCC1 may indicate a favorable outcome in these untreated patients by identifying tumors that have progressed relatively little at the molecular level. Intact DNA repair mechanisms prevent accumulation of genetic aberrations that confer a high malignant potential (45). A recent meta-analysis failed to support two common ERCC1 gene polymorphisms ERCC1 C118T/C8092A and ERCC2 Lys751Gln/Asp312Asn as useful prognostic factors for assessing treatment response to platinum-based chemotherapies in NSCLC patients (46).

Additional biomarkers related to the repair of cisplatin-induced DNA damage have been included in the IALT bio project in order to enhance the predictive power of ERCC1. MutS homolog 2 (MSH2) is a major active component of the mismatch repair machinery; IHC expression of MSH2 displayed a very similar pattern of significance to ERCC1. Specifically, patients with low MSH2 levels had markedly better survival with adjuvant chemotherapy. Those with high levels seemed to have no benefit and in fact had a median survival 9 months shorter than those who did not receive chemotherapy, although this was not statistically significant. Similar to the prognostic role of other DNA-excision-repair proteins, high MSH2 levels predicted significantly longer survival in patients in the observation arm. When MSH2 and ERCC1 expression patterns were combined to form four phenotypes, the benefit from chemotherapy was significantly greater for patients with double-negative tumors. This was also noted when MSH2 expression was combined with that of p27, suggesting that MSH2 immunostaining was a superior predictive biomarker when considered jointly with either of the two other variables (36).

**Ribonucleotide reductase subunit M1 (RRM1)**

RRM1 is a regulatory component of ribonucleoside-
diphosphate reductase, a key enzyme in DNA synthesis that catalyzes the formation of deoxyribonucleotides, by reducing ribonucleotides. The reaction requires generation of a radical allowing the 2'-hydroxyl of ribose to be reduced, which is carried out by the RRMI enzyme (47). The antimetabolite gemcitabine interferes with the function of RRMI by reducing the pool of deoxyribonucleotide-5'-diphosphate available for DNA synthesis (45). Although relevant data for the adjuvant setting are lacking, correlative studies within randomized clinical trials in advanced NSCLC have shown that RRMI overexpression, either at the mRNA or protein level, predicts poor response to gemcitabine-based chemotherapy (48-50). It is somewhat surprising that RRMI protein expression has recently predicted outcome in patients treated with cisplatin and vinorelbine in a biomarker study. In this treatment arm, patients without RRMI protein expression showed improved disease control rates, progression-free survival and overall survival, while RRMI had no predictive impact in patients treated with cisplatin, paclitaxel and gemcitabine (51).

In contrast to predicting chemoresistance, RRMI is a biologically and clinically important determinant of malignant behavior in NSCLC whose overexpression seems to confer favorable outcome. RRMI suppresses cell migration and metastasis, which is at least partially mediated through induction of the phosphatase and tensin homolog gene (PTEN). RRMI is in a region of frequent loss of heterozygosity (LOH), and LOH at this locus was found to be an independent determinant of poor survival in a large cohort of patients with stage I and II NSCLC (52). High RRMI transcriptional expression, defined as mRNA levels above the median value, was favorably prognostic of survival in two independent cohorts of patients with resected NSCLC, most of whom were diagnosed at early stage and treated with surgery only. In this study, RRMI overexpression was a stage-independent predictor of survival, albeit highly correlated with PTEN expression (53). Longer overall survival was recently found in another group of NSCLC patients with high RRMI mRNA expression who had undergone curative lung resection (54). The prognostic role of RRMI was also confirmed by Zheng et al. who measured RNA expression of RRMI and ERCC1 using RT-qPCR in fresh frozen and formalin fixed paraffin-embedded tumor samples (43). This study showed that RRMI expression correlated with ERCC1 expression and that patients whose tumors had high expression of RRMI had superior survival compared with the low expression group (43). In contrast to the previous study, there was no correlation with PTEN expression at the protein level. Interestingly, the concomitant high expression of RRM1 and ERCC1 delineated a subgroup of patients with excellent survival outcomes, accounting for 30% of the cohort (43).

**Breast cancer susceptibility gene 1 (BRCA1)**

BRCA1 is a multifunctional nuclear phosphoprotein which is ubiquitously expressed in all tissues and serves in part as a tumor suppressor, a “caretaker” and a “gatekeeper” in preserving genomic stability. BRCA1 has recently emerged as one of the most appealing biomarkers for personalizing chemotherapy in NSCLC. It has been implicated in normal cellular functions including cell cycle regulation, replication, mitotic spindle assembly, transcription regulation and higher chromatin hierarchical control (55). Also, BRCA1 has a crucial role in DNA repair as a component of the transcription-coupled NER and the homologous recombinant repair pathways. BRCA1 functions as a sensitizer to apoptosis induced by antimicrotubulin agents, such as taxanes and vinca alkaloids and also abrogates apoptosis induced by a range of DNA-damaging agents, including cisplatin and etoposide. Upstream activity of the receptor-associated protein 80 (RAP-80) is required for localization of BRCA1 to sites of DNA double-strand breaks (55).

In a recently reported feasibility study, adjuvant chemotherapy was customized based on BRCA1 mRNA levels in 84 patients with completely resected NSCLC. Patients with higher BRCA1 transcriptional levels were treated with single agent docetaxel, whereas those with intermediate and low BRCA1 expression received cisplatin-based doublets. Interim analyses showed that single-agent docetaxel was not inferior to cisplatin/docetaxel in terms of survival in patients with high BRCA1 levels (56). Therefore, high BRCA1 predicts resistance to cisplatin and possibly sensitivity to docetaxel. Expression levels of BRCA1, divided in quartiles, were assessed in a cohort of 55 patients with stage II to IIIA NSCLC who received neoadjuvant chemotherapy with cisplatin/gemcitabine followed by complete resection. Those with the lowest levels of BRCA1 mRNA expression had significantly greater benefit from chemotherapy in terms of clinical and pathological downsizing as well as overall survival (57).

The potential prognostic role of BRCA1 was investigated in two independent cohorts of chemo naïve patients with early-stage NSCLC analyzed by RT-qPCR.
In the study by Rosell et al., expression level of nine genes involved in DNA repair, including BRCA1, were correlated with overall survival in 126 NSCLC patients who had undergone complete resection and did not receive adjuvant chemotherapy (40). In a univariate analysis, three genes appeared to influence relapse: myeloid zinc finger 1 (MZF1), thioredoxin-1 (TRX1), and BRCA1. However, only BRCA1 and stage III disease remained significant predictors of survival in the multivariate analysis (40). For the 40 patients with a high level of BRCA1 expression, median survival was 29 months while median survival was not reached for the 83 patients with low BRCA1 expression (40). The striking lack of prognostic significance of other biomarkers included in this study may be partially due to the strong intergene coexpression, such as that between BRCA1 and ERCC1, observed. The independent adverse prognostic effect of high BRCA1 expression was confirmed in another cohort of patients with early stage NSCLC also evaluated for ERCC1 and RRM1 mRNA levels. In this study, xeroderma pigmentosum complementation group G (XPG), a key gene for the NER system, was identified as an independent favorable predictor of survival outcome, as well as a potential modulator of recurrence risk among patients with BRCA1 overexpression (58).

Thymidylate synthase (TS)
Thymidylate synthase (TS) and methylenetetrahydrofolate reductase (MTHFR) play important roles in folate metabolism. TS is an enzyme involved in purine synthesis and, as an anti-cancer chemothera-py target, can be inhibited by TS inhibitors such as fluorinated pyrimidine fluorouracil or certain folate analogues, most notably pemetrexed. Consistent findings across phase III trials in advanced NSCLC have established the favorable predictive effect of non-squamous cell histology on treatment with pemetrexed (59). Differentially high TS expression in squamous cell NSCLC represents the main molecular basis underlying this treatment by histology interaction. Data from a current study indicate higher TS expression levels in squamous cell and in high-grade carcinomas (60). No clinical data exist to confirm the predictive role of either histology or TS expression in the adjuvant setting. However an independent prognostic effect for TS has been revealed in chemonaïve patients with resected early stage NSCLC. TS protein expression was correlated significantly with higher proliferative activity of NSCLC cells and, consequently, with poor prognosis in patients with NSCLC who had higher TS level (61). Two other different studies of chemonaïve patients with resected early stage NSCLC revealed an independent prognostic effect for TS, but with conflicting qualitative results. High TS expression at the mRNA, but not IHC level, was significantly associated with adverse disease-free survival in the study from Shintani et al. High TS expression as determined by automated in situ protein quantification, but not by RT-qPCR, predicted improved overall survival in the latter study, in which also TS protein levels did not correlate with those of ERCC1 and RRM1 (62,63). No correlations between intratumoral TS levels and any known clinicopathological variables were reported, with the exception of a recently published article in which TS gene expression was associated with disease stage, lymph node metastasis, tumor differentiation, prognosis, and tumor cell proliferation in patients with lung adenocarcinoma (64).

Kirsten-rous avian sarcoma (KRAS) and p53
The Kirsten-rous avian sarcoma (KRAS) protein is a member of the RAS family of proteins that encode small GTPases involved in cellular signal transduction. Activation of Ras signalling causes cell growth, differentiation, and survival, by transmitting signal downstream from growth factor receptors, including epidermal growth factor receptor (EGFR). From the three RAS genes, KRAS contains > 90% of the mutations detectable in almost 10-25% of NSCLC and mostly affecting codon 12 and 13 (65). Recent data suggested that KRAS mutations may affect outcome of NSCLC patients receiving chemotherapy. In the adjuvant setting, data from the JBR10 trial suggested no benefit from adjuvant chemotherapy in KRAS mutated patients (3,37). Similarly, a retrospective analysis of patients with stage IB disease enrolled in the phase III CALGB-9633 study showed that, among those with tumors larger than 4 cm, KRAS mutations may predict less overall survival benefit from the combination of carboplatin plus paclitaxel (66). In the LACE-BIO pooled analysis, the prognostic and predictive role of KRAS mutations was investigated in 1751 patients treated with adjuvant chemotherapy. Among evaluable patients, KRAS mutations had no effect on survival (67). It should be stressed that the formal test for interaction between the biomarker and treatment effect was not statistically significant in any of the above studies. KRAS mutation status is associated with cigarette smoking and adenocarcinoma histology. The role of KRAS mutations as a prognostic factor in NSCLC remains controversial. Although some studies suggested a potential
negative prognostic effect, other studies did not confirm any negative impact on survival for individuals harboring KRAS mutation (68-71).

The TP53 gene, located on the short arm of human chromosome 17, encodes for a nuclear phosphoprotein involved in the regulation of cell proliferation. The tumor suppressor protein, p53, has a wide range of functions, most of which are mediated via regulation of gene transcription. Commonly described as ‘the genome guardian’, p53 is involved in important cellular processes, such as stress response, cell-cycle control, DNA repair and apoptosis. The mutant gene product, which tends to accumulate to high levels in cancer cells, is believed to exert a dominant negative effect over coexpressed normal TP53. In resected lung cancers, point mutations of the TP53 gene have been found in all histologic types, including approximately 45% of resected NSCLC and, even more frequently, in small-cell lung cancer. Similar to observations with KRAS mutations, p53 mutations have been retrospectively correlated with clinical features such as younger age and squamous histology, but not sex, tumor stage, nodal status, neuroendocrine differentiation, or prior chemotherapy. Many retrospective studies have examined the prognostic role of p53 gene mutations in NSCLC. However, most of these studies have been limited by small size, heterogeneous patient samples, potential selection biases, and/or insensitive p53 mutation detection techniques, leading to inconsistent results. p53 mutations have been associated with decreased survival, no statistically significant change in survival, or improved survival in NSCLC. Previous meta-analyses have indicated that TP53 mutations and p53 expression are weak predictors of outcome in NSCLC (72,73). In contrast, the first published prospective trial examining the prognostic role of p53 mutations in NSCLC demonstrated that neither p53 expression nor TP53 mutations were shown to have predictive value. However, this should be interpreted taking into consideration the different study design and the use of an old-generation regimen (74). However, in the retrospective companion analysis of the phase III NCIC CTG-JBR.10 adjuvant trial p53 IHC overexpression was found to be an independent unfavorable prognostic factor among patients in the observation arm. In addition, only patients with p53-positive tumors derived benefit from the cisplatin plus vinorelbine combination. In contrast to p53 expression, TP53 mutation status was neither prognostic for survival, nor predictive for efficacy of adjuvant chemotherapy (37). This suggests that the biological effects of TP53 mutations and p53 protein overexpression are not equal, highlighting their complex role in tumor aggressiveness and chemosensitivity. Finally, a recent biomarker study as part of the phase III CALGB-9633 adjuvant trial identified p53 and mucin overexpression as independent adverse prognostic factors for stage IB patients (38).

β-TUBULIN and epidermal growth factor receptor

β-tubulin is an essential element of microtubules, which, in turn, serve as cellular structural components involved in vital processes, including mitosis. Among described mechanisms of resistance to anti-tubulin agents, class III β-tubulin (βTUBIII) overexpression is of particular interest. To assess whether βTUBIII might be a useful marker in early NSCLC patients undergoing adjuvant chemotherapy with a vinorelbine-based regimen, levels of βTUBIII were measured in tumor samples from patients treated in the NCIC CTG JBR.10 study. No significant interaction between the biomarker and the effect of cisplatin plus vinorelbine combination was found. Subgroup analysis suggested that high, rather than low, βTUBIII levels were predictive for chemotherapy benefit. However, high βTUBIII expression was shown to be an independent adverse predictor of recurrence-free survival (39). Its prognostic value was retrospectively confirmed in patients enrolled in the IALT study (75). The adverse prognostic significance of high βTUBIII expression is consistent with prior published reports in the setting of advanced NSCLC. Rosell et al. correlated high βTUBIII mRNA levels with inferior outcome in advanced NSCLC patients treated with anti-tubulin agents (76). It has also been shown that high level of expression of βTUBIII in tumor cells, assessed by a semiquantitative IHC assay, was associated with a lower response rate and poor prognosis in advanced NSCLC patients receiving vinorelbine-based chemotherapy (77). In a recent study, high tumor expression of βTUBIII, assessed by IHC in 47 NSCLC patients receiving a paclitaxel-based regimen, was predictive of lower response to therapy and inferior survival (78).

Epidermal growth factor receptor (EGFR) status, defined by mutation analysis or amplification by fluorescent in situ hybridization (FISH), was recently explored in correlation with the results of the phase III NCIC CTG-JBR.10 adjuvant trial. Neither sensitizing mutations nor high gene copy were significantly prognostic in the observation arm. Similarly, although there was a trend toward greater benefit from the cisplatin plus vinorelbine
combination, interaction between the biomarkers and treatment effect was not significant (79). After adjusting for covariates, a recent large, prospective, cohort study of patients with stage I-III adenocarcinoma, 20% of whom had received perioperative chemotherapy, failed to show any significant association between overall survival and mutation status of either EGFR or KRAS (80). In conclusion, the potential prognostic and predictive effect of EGFR amplification and the two most prominent mutations, del 19 and L858R, with regard to chemotherapy effect in the adjuvant setting remains undefined.

**Conclusions**

Finally, are we ready to adopt the use of biomarkers into early stage NSCLC staging, prognosis and treatment selection? Survival amongst cancer patients has improved in recent decades with the availability and application of various treatment modalities. Tumor classification, stage and, sometimes, grade are used to assess prognosis. Although adjuvant chemotherapy has been well established for patients with early stage NSCLC, stage alone is not an ideal biomarker to predict the utility of chemotherapy as the vast majority of patients derive no benefit from treatment. The discovery of molecular biomarkers with the potential to select high-risk patients and predict drug efficacy is essential, especially in controversial fields such as treatment of elderly patients and stage I disease. Biomarker expression often supplants or complements tumor classification, stage or grade. In recent years, a widespread search for new, tumor biology-driven therapeutics has begun, especially in advanced NSCLC. However in the adjuvant setting, it seems that discovery of so-called promising markers translates rather slowly into clinical applicability and few markers have so far been integrated into clinical practice. There are many practical issues, such as the pharmaceutical companies concerns regarding fractionation of markets and medico-legal fears surrounding generation and possession of information. Furthermore, it can take time for physicians and patients to accept and adopt customizing adjuvant chemotherapy. However, as therapies become increasingly target specific, biomarkers will inevitably develop in tandem to play greater roles in staging, grading, and selection of adjuvant therapy; the practical hurdles are many and complex. As already mentioned, the RT-qPCR strategy involving a relatively small number of gene biomarkers and the use of paraffin-embedded specimens seems to outperform wide-genome profiling, although cutoff point definition for continuous variables, such as transcript levels, is particularly challenging because of the great inter-individualization variation of gene expression. In addition, the limited size of most studies and variable techniques used for marker determination plays a role. Often, initially promising results are not reproducible. Another important point is the possible discordance of biomarker status between different types of assays and the corresponding differences in association with clinical outcomes. As already discussed, mRNA expression of a biomarker gene does not necessarily correlate with protein levels as determined by IHC. Apart from multiple technical issues that potentially affect the results of each method, biomarker expression at the protein level depends on additional translational factors, such as microRNA, posttranslational modifications and degradation.

Although the cisplatin plus vinorelbine doublet is currently the standard option for adjuvant chemotherapy, use of appropriate surrogate biomarkers would facilitate randomized clinical trials to establish alternative or superior regimens with smaller sample sizes and shorter follow-up time. The prognostic and/or predictive role of many of the aforementioned biomarkers has been strongly supported by retrospective translational studies. Ideally, biomarkers should be validated analogously in prospective, well-controlled clinical studies of diverse patients across multiple institutions, with well-established standards for all steps of the process. We are anxiously awaiting the results of four prospective multicentre clinical trials of customized adjuvant strategy currently underway (Table 4). ERCC1, RRM1 and BRCA1 are considered to be among the most promising biomarkers with stage-independent, combined prognostic and predictive value, the clinical utility of which is being validated in the ongoing large-scale, randomized phase II and III trials. Until the, highly anticipated, results are in, neither these nor other candidate biomarkers should be used in daily clinical practice as decision-making criteria.

These steps towards personalized medicine will hopefully represent a shift in the management of early staged NSCLC. Indeed, NSCLC should no longer be viewed as one common generic disease, but rather as a collection of rarer tumors with differing biological behaviors and sensitivities to various systemic treatments.

**Acknowledgements**

Disclosure: The authors declare no conflict of interest.
Table 4 Ongoing prospective biomarker studies in early stage non-small cell lung cancer

<table>
<thead>
<tr>
<th>Trial</th>
<th>Stage/histology</th>
<th>Biomarker(s)</th>
<th>Study design</th>
<th>Treatment arms</th>
</tr>
</thead>
<tbody>
<tr>
<td>SWOG-S0720</td>
<td>I (pT ≥2 cm)</td>
<td>ERCC1, RRM1 (AQUA ± RTqPCR ± protein</td>
<td>Phase II</td>
<td>Cisplatin/gemcitabine (↓RRM1 and/or ERCC1) versus observation (↑RRM1 and ERCC1)</td>
</tr>
<tr>
<td>ITACA</td>
<td>II-III</td>
<td>ERCC1, TS (RT-qPCR)</td>
<td>Phase III</td>
<td>Pharmacogenomic-guided chemotherapy regimen versus control*</td>
</tr>
<tr>
<td>TASTE</td>
<td>II-IIIA (pN1)</td>
<td>ERCC1, EGFR (mutations)</td>
<td>Randomized phase II</td>
<td>Customized treatment versus cisplatin/pemetrexed**</td>
</tr>
<tr>
<td>GECP-SCAT</td>
<td>pN1-N2</td>
<td>BRCA1 (RT-qPCR)/EGFR, KRAS (mutations)</td>
<td>Phase III</td>
<td>Customized chemotherapy*** versus cisplatin/docetaxel±RT (pN2)</td>
</tr>
<tr>
<td>RADIANT</td>
<td>IB-IIIA</td>
<td>EGFR (IHC±FISH)/ EGFR, KRAS (mutations)</td>
<td>Phase III</td>
<td>Erlotinib versus placebo ± adjuvant chemotherapy (up to 4 cycles)</td>
</tr>
</tbody>
</table>

AQUA, automated quantitative analysis; SWOG, Southwest Oncology Group; ITACA, International Tailored Chemotherapy Adjuvant; TASTE, Tailored Post-Surgical Therapy in Early Stage NSCLC; GECP-SCAT, Spanish Lung Cancer Group (Grupo Espanol de Cancer de Pulmon)-Spanish Customized Adjuvant Trial; RADIANT, Randomized Double-blind Trial in Adjuvant NSCLC with Tarceva; * taxane (↑ERRC1 and TS mRNA levels), or pemetrexed (↑ERRC1, ↓TS mRNA levels), or cisplatin/gemcitabine (↓ERRC1, ↑TS mRNA levels), or cisplatin/pemetrexed (↑ERRC1 and TS mRNA levels) versus standard chemotherapy with a platinum-based doublet chosen by the investigator; **Customized treatment according to EGFR and ERCC1 status: erlotinib (EGFR-mutant tumors) or observation (EGFR-wild type and ERCC1-positive tumors) or cisplatin/pemetrexed (EGFR-wild type and ERCC1-negative tumors); ***Customized chemotherapy: docetaxel (high BRCA1 levels) or cisplatin/docetaxel (intermediate BRCA1 levels) or cisplatin/gemcitabine (low BRCA1 levels)

References


