Immunohistochemical markers of prognosis in non-small cell lung cancer: a review and proposal for a multiphase approach to marker evaluation

C-Q Zhu, W Shih, C-H Ling,* M-S Tsao

Characteristics of the tumour that affect and predict the survival outcome of patients with cancer are prognostic markers for cancer. In non-small cell lung carcinoma (NSCLC), stage is the main determinant of prognosis and the basis for deciding options for treatment. Patients with early-stage tumour are treated by complete surgical resection, which is curative in 40–70% of patients. That there are other factors important in determining the biology of these tumours, especially genes that have a role in metastasis, is indicated. Such factors could potentially be used to further classify patients into groups according to substages that may be treated differently. During the past decade, a large number of proteins that are putatively important in carcinogenesis and cancer biology have been studied for their prognostic value in NSCLC, but none of them have been proved to be sufficiently useful in clinical diagnosis. Several markers (epidermal growth factor receptor, human epidermal growth factor receptor 2, Ki-67, p53 and Bcl-2) have been studied exhaustively. Ki-67, p53 and Bcl-2 are suggested to be important but weak prognostic markers, by meta-analyses of the results. Cyclin E, vascular endothelial growth factor A, p16INK4A, p27kip1 and β-catenin are promising candidates, but require further study in large randomised clinical trial samples by using standardised assays and scoring systems. Some issues and inconsistencies in the reported studies to date are highlighted and discussed. A guideline for a multi-phase approach for conducting future studies on prognostic immunohistochemistry markers is proposed here.

Lung cancer is the leading cause of cancer death in North America and throughout the world. In North America, annual deaths from lung cancer are greater than the next three most common cancers combined (breast, prostate and colon). Non-small cell lung carcinoma (NSCLC) accounts for about 80% of all lung cancers. The current management of NSCLC is largely guided by tumour stage. Patients with early stage (I and II) tumour are treated by surgical resection with or without adjuvant chemotherapy, and stage III patients require combined modality approaches that may include chemotherapy, radiation and surgery. Nevertheless, the overall 5-year survival rates of these patients remain relatively poor, ranging from 70% for stage IA patients to 25% for stage IIIA patients whose tumours are surgically resectable.1 Most deaths are caused by metastatic recurrence. Differing survival outcomes among patients within a stage suggests the existence of other tumour factors affecting prognosis.

Cancer cells manifest complex genetic aberrations that occur during multi-stage carcinogenesis. Genomic instability or selection leads to aberrations that can be grouped into six essential pathways: the acquisition of (1) self-sufficient or autonomous growth signals; (2) insensitivity to growth-inhibitory signals; (3) resistance to signals of apoptosis; (4) unlimited proliferation potential; (5) sustained angiogenesis; and (6) invasion and metastasis.2 Each of these pathways is regulated by further sets of interacting subpathways, which result in redundancy and additional complexities on the roadmap to malignancy. Despite this, some molecular aberrations are more likely than others to influence the clinical behaviour of a cancer, including the risk of metastasis. Such aberrations, once identified, could potentially serve as prognostic markers, which are tumour (or patient) characteristics that may influence and predict the clinical outcome of a cancer patient.

Molecular prognostic markers could potentially be represented by changes in gene copy number, messenger RNA (mRNA) expression or protein expression levels. Immunohistochemistry (IHC) is the most practical method of assessing protein expression changes in histopathology. IHC not only provides a semiquantitative assessment of protein abundance but also defines the cellular localisation of expression. It may also detect functionally important post-translational protein modifications, such as phosphorylation. These

Abbreviations: ADC, adenocarcinoma; CCN, cyclin; CDK, cyclin-dependent kinase; ECM, extracellular matrix; EGFR, epidermal growth factor receptor; FHit, fragile histidine triad; HER, human epidermal growth factor receptor; HGF, hepatocyte growth factor; hTERT, human telomerase reverse transcriptase; IHC, immunohistochemistry; MMP, matrix metalloproteinase; mRNA, messenger RNA; NSCLC, non-small cell lung carcinoma; PCNA, proliferating cell nuclear antigen; SQCC, squamous cell carcinoma; TGF-β, transforming growth factor β; TIMP, tissue inhibitors of metalloproteinase; VEGF, vascular endothelial growth factor
considerations have led to the extensive use of IHC in studies on prognostic markers for tumours. In this review, we shall summarise to the best of our ability the results of these studies on NSCLC.

MATERIALS AND METHODS
We searched PubMed with the MeSH terms “non-small cell lung carcinoma” and “immunohistochemistry” and “prognosis”, with the search limited to “humans” and “English language”. This search produced 462 papers and 12 reviews dating from May 1987 to October 2005. For individual markers, additional and confirmatory searches were carried out with “gene of interest” and “non-small cell lung cancer” and “immunohistochemistry”. When the same group of investigators published multiple manuscripts on a marker and used overlapping patient cohorts, only the most recent one or the one directly dealing with its prognostic value was emphasised. If meta-analyses or reviews on a specific marker were published, these were considered in place of the original reports (table 1). To limit the scope of this review, we have discussed in more detail only markers that have figured in four or more independent studies. Reports in journals that are difficult to access were also excluded from detailed review. A summary of results from these studies has been compiled and is available at http://www.uhnres.utoronto.ca/labs/tsao/ under supplementary data.

The prognostic significance of a marker is normally evaluated by Cox proportional regression analysis or log-rank testing for its association with disease-free or overall survival. Independent prognostic markers are identified by multivariate analysis, which adjusts for other clinical and pathological factors.

RESULTS
Self-sufficiency for growth pathway
Cells require mitogenic growth signals to enter into the cell cycle. Neoplastic cells invariably show reduced dependence on such signals. This may result from altered expression of soluble growth factors or their receptors, altered cell–cell or cell–stroma interactions and intrinsic activation of growth signal transducers.6 Interactions may be mediated by autocrine, paracrine or endocrine mechanisms.7 In an autocrine loop, cells express both the receptor and its ligand(s). Interactions between tumour and host stromal cells exemplify the paracrine interactions, whereas hormones are the prototype of endocrine signalling.

Epidermal growth factor receptor family
The epidermal growth factor receptor (EGFR)/ErbB family of trans-membrane receptor tyrosine kinases consists of four members: EGFR/human epidermal growth factor receptor-1 (HER-1)/ErbB-1, HER-2/new/ErbB-2, HER-3/ErbB-3 and HER-4/ErbB4.8 These proteins share 40–50% amino-acid homology and have a common domain organisation. On activation by the ligand, these receptors signal through the RAS–RAF–mitogen-activated protein kinase and phosphatidylinositol 3’-kinase pathways, which play critical roles in regulating cell proliferation and survival.

Meert et al9 systematically reviewed 16 studies published between 1989 and 2001, of which 14 were based on immunohistochemistry (IHC). The overall EGFR expression rate in NSCLC was 51%. Expression was less frequent in adenocarcinoma (ADC, 46.2%) than in squamous cell carcinoma (SQCC, 82.6%). Three quarters of the studies (12/16) did not find EGFR expression to be prognostic. A quantitative meta-analysis of eight evaluable IHC studies showed that EGFR IHC expression was marginally significant as a negative (poor) prognostic marker (table 1). In another review, Nicholson et al10 covered reports published during 1985–2000, and found that only 20% and 10% of studies associated increased EGFR levels with decreased recurrence-free survival and overall survival, respectively. Seven additional studies have since confirmed that EGFR IHC is not a prognostic marker for NSCLC,9,11 whereas one study with questionable analytical methods reported a prognostic significance.12 Recent, albeit preliminary, evidence, however, suggests that EGFR IHC may yet assume an important role as a predictive marker for response and survival benefit in patients with advanced NSCLC treated by the small-molecule EGFR inhibitors gefitinib and erlotinib.13 16

A review of 44 reports (39 studies before August 2004) on HER-2 expression in NSCLC showed an overall positive staining rate of 35%, using cut-offs chosen by the various investigators.15 HER-2 overexpression differed significantly between histological subtype (38% in ADC, 16% in SQCC and 18% in large cell carcinoma, p<0.0001). Among 20 studies (including two that used non-IHC assays) with data evaluable for meta-analysis, eight IHC studies reported a significant detrimental effect of HER-2 expression on survival, whereas 10 reported no significant prognostic value. A meta-analysis of these studies suggested that HER-2 overexpression was a relevant poor prognostic marker for 3-year and 5-year survivals, especially among patients with ADC (table 1). The authors, however, warned that several minor studies were excluded from the meta-analysis as they failed to meet the eligibility criteria for meta-analysis; thus the results require caution in interpretation. Two additional studies were published after the meta-analysis. A study on tissue microarray of samples from 284 patients reported that HER-2 expression was a significant predictor of poor survival.

Table 1 Summary of meta-analyses of the results of studies on candidate immunohistochemistry markers for survival of patients with non-small cell lung carcinoma

<table>
<thead>
<tr>
<th>Marker</th>
<th>First author</th>
<th>Published studies</th>
<th>Number of eligible studies</th>
<th>Number of patients</th>
<th>Overall HR*</th>
<th>95% CI</th>
<th>Significance/comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGFR</td>
<td>Meert</td>
<td>Until July 2001</td>
<td>8</td>
<td>1987</td>
<td>1.13</td>
<td>1.00 to 1.28</td>
<td>Weak significance</td>
</tr>
<tr>
<td>P21RAS</td>
<td>Mascoux</td>
<td>Until July 2003</td>
<td>7</td>
<td>989</td>
<td>1.08</td>
<td>0.86 to 1.34</td>
<td>Not significant</td>
</tr>
<tr>
<td>HER-2</td>
<td>Nakamura</td>
<td>Until August 2004</td>
<td>18</td>
<td>2579</td>
<td>1.32</td>
<td>1.14 to 1.55</td>
<td>Weak significance</td>
</tr>
<tr>
<td>PS3</td>
<td>Steels</td>
<td>Until July 1999</td>
<td>8 (Pab1801)</td>
<td>1035</td>
<td>1.57</td>
<td>1.28 to 1.91</td>
<td>True significance requires prospective and multivariate confirmation by a standardised technique, scoring criteria and cut-offs</td>
</tr>
<tr>
<td>Ki-67</td>
<td>Martin</td>
<td>Until December 2002</td>
<td>16</td>
<td>2067</td>
<td>1.25</td>
<td>1.09 to 1.43</td>
<td>False significance</td>
</tr>
<tr>
<td>Bcl-2</td>
<td>Martin</td>
<td>1993 to December 1999</td>
<td>18</td>
<td>2909</td>
<td>0.72</td>
<td>0.64 to 0.82</td>
<td>False significance</td>
</tr>
</tbody>
</table>

CI, confidence interval; EGFR, epidermal growth factor receptor; HER, human epidermal growth factor receptor; HR, hazard ratio of death for high expression of the marker.

*HR<1 implies worse survival for the group with increased expression of the marker, whereas HR>1 (Bcl-2) indicates better survival for the group with higher expression. For PS3, the prognostic significance of studies that used Pab1801 and DO7 antibodies were separately analysed.
in patients with lung ADC, but only 2 of 80 patients expressed this protein. A second study on 345 patients also included a rather extensive comparative analysis of various scoring criteria, but reported no noteworthy association between HER-2 expression and prognosis. Overall, the prognostic significance of HER-2 has not been established.

**Hepatocyte growth factor and Met receptor**

Hepatocyte growth factor (HGF) is also known as scatter factor, a multifunctional cytokine with putative roles in cell proliferation, motility, angiogenesis and morphogenesis. Met is the tyrosine kinase receptor for HGF/scatter factor. The prevalent concept regarding HGF-Met signalling is that it is paracrine, with the Met receptor expressed by epithelial cells and the HGF ligand secreted by stroma fibroblasts. Tsao et al. showed common expression of HGF mRNA in lung adenocarcinoma (ADC) and protein in tumour cells, indicating that HGF-Met autocrine signalling also plays an important role in NSCLC. The reported rates of Met expression or overexpression in NSCLC were 40–60%.

Takanami et al. reported that Met expression was an independent poor prognostic marker in ADC. In contrast with other studies that did not find considerable Met expression in tumour stroma cells, Tokunou reported that Met stromal overexpression in 53% of lung ADC, and this staining was associated with shorter survival for the patients. Takanami et al. also reported that tumour expression of HGF in ADC was a poor prognostic marker, but only in univariate analysis. Overall, the prognostic roles of Met and HGF in NSCLC remain uncertain, and additional studies using well-characterised and specific antibodies are warranted.

**Growth signal transducers**

RAS is a small G-protein with intrinsic GTPase activity. It is a critical regulator of signalling downstream of cell surface receptors. Point mutations on codon 12, 13 or 61 of the ras family genes result in their oncogenic activation, whose significance in NSCLC has recently been proved by a transgenic mouse model. Ras mutations are found in approximately 20% of NSCLC, and they mainly occur in ADC or large-cell carcinoma. The prognostic significance of ras mutation remains indeterminate, but the frequency and significance of RAS protein overexpression in lung cancer has been poorly investigated. A meta-analysis of seven evaluable IHC studies showed the absence of prognostic value for RAS protein expression in NSCLC. One possible lack of enthusiasm to study RAS protein expression could be the lack of antibodies with confirmed specificity and reactivity for use in IHC.

**Cell cycle checkpoint promoters**

Cell cycle includes the G0 (quiescent), G1 (resting), S (DNA synthesis), G2 (pre-mitotic gap) and M (mitotic) phases. Progression between these phases is an orderly process tightly regulated by complex yet redundant mechanisms that include multiple checkpoints, which are used to assess growth signals, cell size and DNA integrity. Cell cycle progression is promoted by cyclins (CCNs) and CCN-dependent kinases (CDKs). The constitutive and persistent high expression of these regulators may cause failure of checkpoint arrests and may lead to uncontrolled proliferation. The CCN family has 18 members (A–I, L and T), each of which may have multiple isoforms (eg, CCND1–3). The CDKs are also composed of several family members, from CDK1 (cdc2) to CDK11. The key CCNs and CDKs are CCNA (S and G2 phases), CCNB (G2/M transition), CCND and CCNE (G1/S transition), as well as CDK1 (G2/M transition), CDK2 (S phase), CDK4 and CDK6 (G1/S transition). CCND1 is the most studied CCN in NSCLC. Among 15 reports reviewed, CCND1 expression or overexpression was noted in approximately 50% of NSCLC, but its overall prognostic value remains uncertain. Five studies identified CCND1 overexpression as a negative prognostic marker, whereas three others associated it with better prognosis. Only two poor prognostic reports remained significant in multivariate analyses. The remaining seven reports reported no association.

The overexpression of CCNE is also common (~45%) in NSCLC. Five of eight studies showed a significant association between CCNE overexpression and worse prognosis, and three of them found it to be a significant independent marker after adjusting for stage. By using expression transcriptions-polymerase chain reaction assay, Muller-Tidow et al. also reported that CCNE mRNA overexpression was an independent poor prognostic marker in 70 patients with NSCLC.

Proliferating cell nuclear antigen (PCNA) is the subunit of DNA polymerase δ responsible for its proofreading activity, which has a marked role in maintaining the fidelity of mammalian DNA replication. Thus, overexpression of PCNA may potentially influence cell cycle progression and modulate the effectiveness of radiation and chemotherapy. At least 12 studies have evaluated the prognostic value of PCNA overexpression in NSCLC, with more studies reporting no prognostic value of PCNA overexpression. KI-67, another nuclear antigen that is expressed only in proliferating cells, preferentially during the late G1, S, G2 and M phases of the cell cycle, is commonly used as a marker to evaluate proliferation of tumour cells. Almost all studies have used the MIB1 antibody to evaluate KI-67 by IHC. Martin et al. recently conducted a meta-analysis of studies published before 2003 and identified only 15 of 37 studies (41%) that reported a negative effect of KI-67 overexpression on prognosis for patients with NSCLC. In all, 16 studies on 1863 patients with NSCLC were evaluable for the meta-analysis. The aggregated survival data showed that KI-67 immunoactivity was associated with poorer survival (hazard ratio (HR) 1.55, 95% confidence interval (CI) 1.34 to 1.78) and the HR remained significant even within histological subtypes. Four of nine additional studies reported since 2002 also reported that KI-67 expression was associated with a poorer prognosis.

**Insensitivity to growth inhibition pathway**

A balance between stimulators and inhibitors of cell proliferation maintains growth homeostasis in normal cells. Resistance to growth-inhibitory factors is an essential step in carcinogenesis. Inhibitory signals may originate from the tumour cells themselves or from their microenvironment. These factors could be soluble molecules, the actions of which are mediated by autocrine, paracrine or endocrine mechanisms. Transforming growth factor β (TGF-β) is generally an antiproliferative factor for epithelial cells and is produced mainly by the stromal cells. Evidence, however, suggests that TGF-β may promote tumour progression. Few studies have explored the prognostic significance of TGF-β expression in NSCLC. These studies have provided contradictory results in both the prognostic significance and cellular localisation of TGF-β expression. CCN-dependent kinase inhibitors are negative regulators of CCNs and CDKs. Two major groups exist: the CIP/KIP family genes (p21, p27 and p57), which inhibit all CDKs, and the INK4 family genes (p16, p15, p18 and p19) that inhibit CDK4/6. The loss of CCN-dependent kinase inhibitors should theoretically, therefore, lead to restraint on the proliferative stimuli of tumour cells, thereby leading to tumour progression and poorer prognosis. Less than half the studies on the prognostic effect of p21WAF1/CIP1 in NSCLC have reported significant differences, although this may be owing to the
noting that unlike many proteins whose protein levels are
been reported in approximately 30% of NSCLC. It is worth

studies that have assessed the clinical effect of p16 protein
inactivated mainly by promoter hypermethylation. Among 12

Retinoblastoma protein, thus releasing the E2F to exert its transcrip-
tional activity. P16

30% of tumours, inactivation is by homozygotic deletion,
whereas inactivation by mutation is rare.75–77 The others are
inactivated mainly by promoter hypermethylation. Among 12

two additional studies found p16 loss to be a poor prognostic marker, but only in early-stage tumours
or SQCC.40–44 Thus, despite some discordance, p16 represents
a promising marker for further evaluation.

Retinoblastoma protein inactivation may achieve the same
results as p16 inactivation and serves as an alternate pathway
to promote G1-S cell cycle progression. Loss of retinoblas-
toma protein expression at variable levels has been reported
in unselected tumours of NSCLC,10–12,124–125 but there is no convincing evidence to suggest that it is a significant
prognostic marker.14–16,41,68–71,91–93,95–97

Resistance to the apoptosis pathway
Apoptosis is orchestrated by the sequential activation of
caspases, a family of cysteine proteases with specificity for
aspartic acid residues. Two pathways may mediate apoptosis:
(a) the death receptors pathway (tumour necrosis factor, Fas
ligand or TRAIL receptors) that activates caspase 8 and 10
and (b) the mitochondrial pathway that activates caspase 9.
Both pathways lead to effector caspases (3, 6 and 7).128 Bcl-2
is antiapoptotic for the mitochondria pathway, whereas Bax
is a proapoptotic factor activated by p53.

p53 gene is the most studied gene in all types of cancer,
including lung cancer. IHC can detect missense mutant p53
proteins, but may not detect the protein products of nonsense,
deletion or truncation mutants. This is partially
responsible for the incomplete concordance between p53-
positive IHC and sequencing results.115–126 At least three meta-
analyses have been carried out on the prognostic value of p53
gene alterations in NSCLC.100–102 The frequency of p53
mutation or positive IHC staining is consistently higher in
SQCC than in ADC. Although more studies have reported no
prognostic significance of p53 IHC than studies that
associated it with poor prognosis, the overall conclusion of
these meta-analyses is that p53 staining has some adverse
prognostic effect in patients with NSCLC (table 1). The
strength of this significance, however, remains to be
confirmed prospectively and in clinical trials on large patient
cohorts with more homogeneous and well-defined patient
populations. Since these meta-analyses, more than 20
additional retrospective small cohort studies have been reported.
Unfortunately, overall, they provided similarly
contradictory results as those reviewed in the meta-
analyses.8,19,40–45,65,75,77,85,91–119

MDM2 is the E3 ubiquitin ligase that targets p53 for
nuclear export and degradation. MDM2 and p53 form a
feedback loop that is also modulated by p14ARF, the alternate
spliced product of INK4 gene that also produces p16INKA.
P14ARF binds MDM2, thus the overexpression of p14 would
lead to a sequestration q MDM2 and prevents its binding
with p53, thus resulting in stabilisation of the p53 protein.
This is the postulated mechanism wherein p53 protein
overexpression is found in the absence of mutation.129
The precise role of MDM2 in tumour progression, however,
remains uncertain.129 Although MDM2 gene amplification
has been associated with poor prognosis,129 MDM2 mRNA
expression has also been reported to be a favourable
prognostic marker in patients with NSCLC.130 Five studies
have evaluated the prognostic value of MDM2 IHC in patients
with NSCLC.123–127 Whereas three studies did not find MDM2
expression to be prognostic, two groups have reported MDM2
staining to be a favourable prognostic marker in p53-negative
patients with NSCLC.124–127

At least 20 members in the Bcl-2 family share at least one
conserved Bcl-2 homology domain.126 These are subdivided
into prosurvival (antiapoptosis) family members, including
Bcl-2, Bcl-x, and Bcl-w, and proapoptosis families, including
Bax, Bak and Bok. There are also Bcl-2 homology 3-only
proteins that promote cell death, including Bid, Bim, Bik,
Bad, Bmf, Hrk, Noxa and Puma. The balance between
prosurvival and proapoptosis factors determines whether a
cell will respond to apoptotic signals. A meta-analysis of
publications from 1993 to 1999 examined the prognostic
significance of Bcl-2 in lung cancer (table 1).129 Among 21
studies on NSCLC, 10 reported Bcl-2 expression as a good
prognostic factor, whereas one reported it as a poor
prognostic marker. The expression of Bcl-2 was found in
approximately 35% of NSCLC, being more common in ADC
(61%) than in SQCC (32%). Counter-intuitive to the
antiapoptotic function of Bcl-2, a meta-analysis of 18 studies
with evaluable data and comprising 2909 patients showed a
better survival outcome for patients with Bcl-2-positive
tumours (HR 0.72, 95% CI 0.64 to 0.82). The results were
similar for the two antibodies most commonly used (clone
100 or 124) and the four ranges of thresholds used as cut-off
(table 1). Since the meta-analysis, there have been 27
additional independent IHC studies,44–46,60,101,132–137,119,130–145
most of them using the clone 100 and 124 antibodies. The
mean Bcl-2 expression rate (34%) was remarkably similar to
those included in the meta-analyses, and approximately 40% of
the studies reported Bcl-2 expression to be a good
prognostic marker.

No prognostic significance seems to exist for Bax in
NSCLC,67,92,99,110,140,146–148 Only one of seven studies indicated
an association of Bax expression with better survival,140
whereas the others failed to show significance, including
one study that evaluated only patients treated by chemother-
apy.147

Unlimited growth potential pathway
Telomerase complex, which contains an RNA subunit human
telomerase RNA gene (hTERC) and a protein catalytic
subunit human telomerase reverse transcriptase (hTERT), is
responsible for the maintenance of integrity of chromosomal
telomeres and genomic stability. IHC detection of hTERT
expression is frequent in NSCLC (50–90%)150–152 and has been
correlated with increased telomerase activity,150–152 but it is
not predictive of survival. The specificity and sensitivity of hTERT
antibody for IHC has not proved to be optimal. Even with
the best available monoclonal antibody 44F12 that gives a single
band in western blot, the concordance rates of IHC staining

www.jclinpath.com
and western blot, telomerase activity and mRNA expression as detected by in situ hybridisation were 76%, 70% and 70%, respectively.152

**Sustained angiogenesis pathway**

Angiogenesis is crucial for primary tumour growth and metastasis, and comprises interactions between tumour cells, endothelial cells and stroma cells. Growth factors and cytokines have been implicated in angiogenesis, as well as proteases that break down extracellular matrix (ECM). Vascular endothelial growth factor (VEGF), also known as vascular permeability factor, is a key regulator of angiogenesis in physiological and pathological conditions.153 To date, seven family members of VEGF have been identified, including VEGF-A (also known as VEGF), VEGF-B, VEGF-C, VEGF-D, VEGF-E, placenta growth factor 1 and placenta growth factor 2.154 Overexpression of VEGF has been associated with tumour progression and poor prognosis in most cancer types, and novel antiangiogenic agents have recently been approved or are in various phases of clinical trials for cancer treatment.155 VEGF expression or overexpression has been reported in approximately 60% of NSCLC. Most studies have indicated a significant positive correlation between VEGF immunoreactivity and tumour vascularity.67 156–159 To date, seven of 12 studies have reported a correlation between VEGF immunoreactivity and tumour cell expression.67 150 154–156 160–164 Four of five studies have also shown that VEGF-C is a predictor of poor prognosis, but only one found it to be an independent marker.67 164 167–169

**Invasion and the metastasis pathway**

Aside from angiogenesis, the breakdown of intercellular junctional complexes and ECM proteins by tumour cells is also an important process in tumour progression. Integrons, cadherins, selectins, the immunoglobulin super gene family (IgSF) and CD44 are families of adhesion or junctional molecules with different genetic and biochemical properties. Factors that alter the abundance or function of ECM proteins and adhesion or junctional molecules may influence the ability of tumour cells to invade and metastasise, and thus cancer prognosis.

E-cadherin and catenins are components of adherens junction protein, with crucial roles in the maintenance of intercellular junctions in epithelial cells. Thus, reduced catenin and E-cadherin expression could potentially affect tumour differentiation, metastasis and prognosis. In all, three family members of catenin exist, α, β and γ. Whereas α-catenin and γ-catenin strictly participate in junctional complex formations, β-catenin is also associated with the regulation of transcription through the wingless-type MMTV integration site family pathway.160 161 Expression of catenins has usually been assessed in conjunction with that of their partner protein E-cadherin. Eight studies have evaluated the prognostic role of β-catenin in NSCLC172–179; three of these simultaneously investigated the significance of α-catenins and γ-catenins.172 173 176 Reduced expression of β-catenin was found in about 30% of NSCLC, and five studies reported the association of reduced staining with poor prognosis.172 173 176 177 Interestingly, two of three studies that investigated the α-catenins and γ-catenins also found that their reduced expression was a poor prognostic marker.176 177 Four of seven studies reviewed also reported favourable prognosis for NSCLC tumours that retained normal E-cadherin expression, with two studies remaining significant in multivariate analysis.173 175–177 179–181 Thus, β-catenin and E-cadherin are worthy candidates for further investigation.

Matrix metalloproteinases (MMPs) comprise of almost 20 family members and are the only enzymes that can degrade fibrillar collagens. There are also four tissue inhibitors of metalloproteinase (TIMPs), which are endogenous anti-MMP molecules.182 Although the expression of many of these MMPs and TIMPs has been studied by IHC, most have focused on the gelatinases MMP-2 and MMP-9.163 164 163–168 Results from most of these studies have suggested that high expressions of MMP-2 and MMP-9 are poor prognostic markers for NSCLC, especially the expression of MMP-2. Most of these studies have identified tumour cell expression of the MMPs as prognostic,164 165 166 167 whereas Ishikawa et al183 reported that the stroma, but not tumour cell expression of MMP-2, was prognostic. Increasing evidence resulting from in situ hybridisation and laser capture microdissection coupled with real-time reverse transcription-polymerase chain reaction technique suggests that most MMPs and TIMPs are predominantly expressed in the stroma rather than in the tumour cells,167 whereas in nearly all published studies the focus has been almost exclusively on tumour cell expression, with few studies reporting on both tumour and stroma cells. Overall, the role of MMP IHC in the prognosis of NSCLC should be studied further, but these need to be accompanied by a very rigorous validation of the specificity of the antibodies.

**Other candidate markers**

Our discussion is restricted to markers that have been studied in four or more independent studies. A large number of markers exist, however, which have been less studied but have shown either promising or controversial results as prognostic indicators (fig 1). Markers in which expression changes have been associated with worse prognosis in at least two studies, and which have no reports of an opposite association, are considered to be promising candidate markers. These include CCNA (two studies),169 170 CCNB1 (two of three studies),168 169 carbonic anhydrase IX (two studies)172 173 and fragile histidine triad (FHIT; two of four studies).174 175–177 In these markers positive expression is associated with poorer prognosis, although retained (positive) FHIT expression is associated with better outcome.

**DISCUSSION AND CONCLUSION**

Of 17 markers that have been investigated by eight or more groups (fig 1), none have shown consistent results in all studies. There are, however, six markers (overexpression of CCNE and VEGF-A, and loss of p16INK4A, p27kip1, β-catenin, and E-cadherin) that have shown correlation with poor prognosis in 50% or more of the studies, without contrary results that have been studied most exhaustively (EGFR, HER-2, Ki-67, p53 and Bcl-2) have all reported inconsistent results, perhaps suggesting that their prognostic values are at best weak. In addition to the above markers, there are other promising markers that have been less studied, yet have shown significant results more often than non-significant results. These may also be worthy of further validation, and include HGF/Met, CCNA, CCNB1, VEGF-C, carbonic anhydrase (CA) IX, α-catenin and γ-catenin, MMP-2, MMP-9 and FHIT.

Our review has largely confirmed what is well known among pathologists and oncologists—that studies on prognostic markers, particularly, but not exclusively, those including immunohistochemical assays, often give rise to inconsistent or contradictory results. The issue has become especially urgent as oncology is entering the dawn of personalised medicine with targeted treatments. Recently, the Statistics Subcommittee of the NCI-EORTC Working Group on Cancer Diagnostics has published the REporting recommendations for tumour MARKer prognostic studies (REMARK) guidelines.186 The data we collected during this review (http://www.uhnres.utoronto.ca/labs/
Figure 1  Number of reports for markers that have been studied by two or more independent groups of researchers. For each marker, the graphs also indicate the relative distribution of studies that showed significant association of high expression or overexpression with good (favourable) prognosis, poor prognosis or no prognostic value. The broken line is to facilitate identifying markers that have been studied by eight or more groups. The detailed summaries of markers and their deregulations are available at http://www.uhnres.utoronto.ca/labs/tsao/index.php.
tsao/supplementary-data.php) clearly highlight the multiple study design factors that have contributed to the inconsistencies and contradictions. Many of these issues require serious discussion by the pathologist community.

For almost all markers, diverse sources and types of antibodies have been used. Multiple antibodies per se should not become an issue as long as each antibody has been well characterised for its specificity and sensitivity. For many markers, especially new ones, the quality assurance for the antibody is often not adequately described. Specification of the method used to determine the antibody titre that is finally adopted is often neglected. Including representative low and high photomicrographs of the staining would also go a long way in convincing readers that the staining is probably specific.

It should be recognised that IHC scoring is at best a semiquantitative exercise for which no standard criteria have as yet been proposed or adopted. The heterogeneity in scoring methods possibly partly reflects our difficulty in coming to grips with genotypic and phenotypic (expression) heterogeneity in tumours. Unfortunately, our understanding of the cause and effect of these heterogeneities, including immunostaining, is almost non-existent, thus making it difficult to design rational and biologically meaningful scoring criteria. The scoring methods that have been used included estimating the grade of staining intensity (numerically from absent to strong), percentage of tumour cells stained, cellular localisation of the antigen and systems that combined these parameters. The approaches to combining the intensity and percentage of cells stained have also been divergent, including a direct multiplication of the percentage of stained cells with staining intensity score, or of grouping the staining percentage into 3–4 scores and then summing this with the intensity score. Although superficially the two methods could lead to comparable final scores, in reality, the former could unreasonably overweight the percentage score. The only scoring system that has been generally accepted is the one using the method used to determine the antibody titre that is finally adopted is often neglected. Including representative low and high photomicrographs of the staining would also go a long way in convincing readers that the staining is probably specific.

It should be recognised that IHC scoring is at best a semiquantitative exercise for which no standard criteria have as yet been proposed or adopted. The heterogeneity in scoring methods possibly partly reflects our difficulty in coming to grips with genotypic and phenotypic (expression) heterogeneity in tumours. Unfortunately, our understanding of the cause and effect of these heterogeneities, including immunostaining, is almost non-existent, thus making it difficult to design rational and biologically meaningful scoring criteria. The scoring methods that have been used included estimating the grade of staining intensity (numerically from absent to strong), percentage of tumour cells stained, cellular localisation of the antigen and systems that combined these parameters. The approaches to combining the intensity and percentage of cells stained have also been divergent, including a direct multiplication of the percentage of stained cells with staining intensity score, or of grouping the staining percentage into 3–4 scores and then summing this with the intensity score. Although superficially the two methods could lead to comparable final scores, in reality, the former could unreasonably overweight the percentage score. The only scoring system that has been generally accepted is the one used to assess HER-2/c-erbB2 staining. Perhaps a more important confusion that permeates many biomarker correlative studies, especially IHC, is the method of determining cut-offs for dichotomising scores in log-rank tests. Cut-offs were often arbitrary and sometimes selected to obtain the desired effect by the “minimum p value” approach. Such approaches have led to the existence of diverse but poorly justified cut-offs, which may contribute to non-reproductibility of the results. In general, taking the median as cut-offs for log-rank tests provides the most unbiased, albeit stringent, criteria for analysis, but selecting multiple rationally justified cut-offs is also worthy of performing.

Unlike predictive markers for the efficacy of specific treatment, prognostic markers in NSCLC are most relevant for early-stage or resectable tumours. By including patients with advanced-stage tumour into cohorts of patients with primarily early-stage tumour, we may introduce unwanted effects. Some studies have reported that specific markers could have different prognostic significance in patients with early-stage tumours as compared with those with advanced-stage tumours, or in different histological subtypes.

Unfortunately, subgroup analysis will reduce the number of patients in each subgroup, and thus may underpower the statistical analysis. Biased selection of patient cohorts may also lead to results that will be difficult to reproduce in other studies.

The most relevant end point for assessing the prognostic value of a marker is its association with overall or disease-free survival. Correlations with stage or tumour grade may provide insights into the tumour biology, but are weak and inadequate parameters for assessing the significance of the outcome. As marker expression can be correlated with other clinical and pathological parameters (eg, age, stage, sex), which may also influence prognosis, only multivariate analysis that adjusts for these factors is more likely to become clinically useful. Studies with small sample size (<100 cases) more often than not will fail such stringent statistical tests.

Last but not least, we have limited our review to the prognostic significance of individual markers, although most of them were published in the context of studies on multiple

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Proposed multiphase evaluations of immunohistochemical prognostic markers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>mPhase</strong></td>
<td><strong>Objectives</strong></td>
</tr>
<tr>
<td>1</td>
<td>To verify:</td>
</tr>
<tr>
<td></td>
<td>1. Specificity of antibody and assay technique</td>
</tr>
<tr>
<td></td>
<td>2. Cellular localisation of expression</td>
</tr>
<tr>
<td></td>
<td>3. Expression changes in tumour cells compared with normal cells</td>
</tr>
<tr>
<td>2</td>
<td>To investigate the prognostic significance of markers in samples from a single institution</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>To validate the significance of best candidate markers in phase III clinical trials with tumour samples of limited availability</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>To test prospectively the performance of markers in phase III randomised clinical trials by incorporating markers as stratification factors</td>
</tr>
</tbody>
</table>

CI, confidence interval; HR, hazard ratio.
markers. Many of these studies have reported that whereas an individual marker was not prognostic, a combination phenotype could be prognostic, but such analyses could raise many issues and caveats that are beyond the scope of the current review.

In accordance with the above caveats, we suggest that perhaps studies on tumour biomarkers should adopt a similar multi-phase approach as that being used in clinical trials for new therapeutics (table 2). Most studies may combine marker Phase (mPhase) 1 and 2 studies but even in this early stage of marker evaluation, a conceptual multi-phase approach would provide a more consistent reporting of the early stage of marker evaluation. A conceptual multi-phase marker Phase (mPhase) 1 and 2 studies but even in this multi-phase approach as that being used in clinical trials for perhaps studies on tumour biomarkers should adopt a similar current review.

many issues and caveats that are beyond the scope of the

ACKNOWLEDGEMENTS
This work is supported by grant 02-MAY-0132 from the Ontario Cancer Research Network and grant # 015 184 from the Canadian Cancer Society and National Cancer Institute of Canada. Dr Tsao is the M Qasim Choksi Chairman in Lung Cancer Translational Research. Dr Ling is supported by a visiting scholarship from the China Scholarship Foundation. We thank Dr Sarit Aviel-Ronen and Dr Frances Shepherd for critical comments and Davina Lau and Daniel Gomez for assistance in data collection.

Authors’ affiliations
C-Q Zhu, W Shih, C-H Ling, M-S Tsao, Department of Pathology, University Health Network, Princess Margaret Hospital, Toronto, Ontario, MSG 2C4, Canada

Competing interests: None.

*Current address: The First Affiliated Hospital of Soochow University, Suzhou, China

REFERENCES


120 Midgley CA, Lane DP. p53 protein stability in tumor cells is not determined by mutation but is dependent on Mdm2 binding. Oncogene 1997;15:1179–89.


Immunohistochemical markers of prognosis in non-small cell lung cancer: a review and proposal for a multiphase approach to marker evaluation
C-Q Zhu, W Shih, C-H Ling and M-S Tsao

J Clin Pathol 2006 59: 790-800
doi: 10.1136/jcp.2005.031351

Updated information and services can be found at:
http://jcp.bmj.com/content/59/8/790

These include:

References
This article cites 200 articles, 50 of which you can access for free at:
http://jcp.bmj.com/content/59/8/790#BIBL

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/