Introduction

Molecular testing of targetable alterations in non-small cell lung cancer (NSCLC) is now very well-established in routine practice prior to the prescription of targeted agents. However, in recent years, this has been supplemented by a variety of biomarkers which predict response to immunotherapeutic agents.

Two factors conspire to make testing for these biomarkers challenging. The first is that—particularly in the case of lung—very tiny biopsy samples must be used to assess for the presence of a multitude of molecular targets at the levels of genes, chromosomes and proteins. The second is that prediction of response to immunotherapy is extremely complex: it is clear that no single biomarker will be sufficient to predict response to these agents and, even amongst those which have demonstrable value, methods of assessment and thresholds for positivity remain tentative.

Here, we discuss the current landscape of predictive markers for immunotherapy in NSCLC, and review the challenges presented by them for laboratories providing this service.
PD-L1 expression

Programmed death-ligand 1 (PD-L1) is a cell surface protein expressed physiologically in a variety of tissues. It binds programmed cell death protein 1 (PD-1) on cytotoxic T-cells, inducing anergy and/or apoptosis in the latter and so preventing immune mediated-destruction of the cell (1). By virtue of its downregulatory activity, it is regarded as an “immune checkpoint”, and is thought to play a role in establishing immunological tolerance under normal circumstances (2). This inhibitory function is also germane in the case of cancer cells, whose genetic aberrations lead to the expression of numerous neoantigens which are normally recognised by the immune system, triggering an immediate-mediated anti-tumour response. A proportion of tumours are known to aberrantly express PD-L1, rendering them invisible to the immune system and permitting unchecked tumour growth.

In 2010, a phase I trial demonstrated that immune checkpoint inhibitors showed efficacy in a range of solid tumours, including NSCLC (3). Subsequent studies indicated that cell surface expression of PD-L1 was an effective biomarker in predicting response to these drugs. The Keynote-001 study showed that, when the proportion of tumour cells expressing membranous PD-L1 was assessed by immunohistochemistry, high levels of expression (at least 50% of tumour cells) predicted powerful responses to the anti-PD-1 agent pembrolizumab (4). Later trials compared pembrolizumab with standard chemotherapy, and demonstrated superior efficacy for pembrolizumab in tumours strongly expressing PD-L1 (5,6). Favourable responses have been reported for other checkpoint inhibitors, namely for nivolumab in the Checkmate trials (7,8), and for atezolizumab (9,10).

PD-L1 expression assessment: the assays

The development of checkpoint inhibitor therapy has been very much trial-driven, with each series of trials employing a particular assay to measure PD-L1 expression. The result of this is that the clinical evidence for efficacy of each assay is established in the context of a particular checkpoint inhibitor; prescription of a given drug therefore requires assessment of PD-L1 expression by its companion assay. Four commercial antibodies are currently available to assess PD-L1 expression, using formalin-fixed, paraffin-embedded tissue sections. Each assay employs different means of assessment and thresholds for positivity, which reflect the methodologies employed in their respective trials (Table 1).

This situation is not ideal, and has led to considerable confusion (11). In response, there have been considerable efforts to harmonise the four antibodies, with the hope that they might be used interchangeably for the prescription of any of the anti-PD-1/PD-L1 agents. Early evidence from the Blueprint Project suggests that this may be a possibility; studies have demonstrated high concordance between the 22C3, 28-8 and SP263 assays, potentially paving the way to a degree of interchangeability (12); it should be borne in mind, though, that concordance between these assays is not perfect, and so a minority of patients’ risks being denied potentially valuable immunotherapy or subjected to unhelpful therapy if the trial assay is substituted for another.

In addition to these companion diagnostics, a mixture of cost pressures and strategic decisions have led many services to make use of laboratory-developed tests (LDTs). These are assays which have been developed in-house and technically validated against the commercial assays. In spite of this, however, such LTDs lack the clinical validation which is provided for the commercial assays by clinical trials.

It is fair to say, then, that at present there is no single “best” means of assessing PD-L1 expression. The situation is confusing, and the question as to how best to perform testing is likely to change as evidence accrues.

PD-L1 expression assessment: quantification

For three of the four companion assays, PD-L1 expression is

Table 1 Comparison of the commercially-available PD-L1 assays

<table>
<thead>
<tr>
<th>Assay</th>
<th>PD-L1 IHC 22C3 pharmDx</th>
<th>PD-L1 IHC 28-8 pharmDx</th>
<th>SP263</th>
<th>SP142</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manufacturer</td>
<td>Agilent</td>
<td>Agilent</td>
<td>Ventana</td>
<td>Ventana</td>
</tr>
<tr>
<td>Assessed expression</td>
<td>Tumour cells</td>
<td>Tumour cells</td>
<td>Tumour cells</td>
<td>Tumour and inflammatory cells</td>
</tr>
<tr>
<td>Companion drug</td>
<td>Pembrolizumab</td>
<td>Nivolumab</td>
<td>Durvalumab</td>
<td>Atezolizumab</td>
</tr>
<tr>
<td>Drug manufacturer</td>
<td>Merck Sharp and Dohme</td>
<td>Bristol-Myers Squibb</td>
<td>AstraZeneca</td>
<td>Roche</td>
</tr>
</tbody>
</table>

PD-L1, programmed death-ligand 1; IHC, immunohistochemistry.
measured in terms of a tumour proportion score (TPS): this is the ratio of tumour cells expressing PD-L1, to those not expressing PD-L1 (Figure 1). The definition of expression is any degree of membranous staining in a tumour cell, of any degree of completeness or intensity, excluding necrotic cells. The SP142 assay differs in that it also incorporates staining in tumour-associated inflammatory cells, in line with the methodology employed in the trials for atezolizumab (13).

It is not surprising, given that the extent of tumour PD-L1 expression predicts response to checkpoint inhibitors, that prescription of these drugs is conditional on the degree of PD-L1 expression. This varies between countries and
 territories in three respects, chiefly as a result of differences in the mechanisms of funding both testing and drug prescription:

(I) Whether the PD-L1 expression status assessment is mandatory prior to prescription, or rather serves as a complementary piece of information to be taken into account before administration of the drug;

(II) Precisely which thresholds are to be considered “positive” for PD-L1 expression;

(III) Closely related to the point above, whether a particular positivity threshold should be deemed acceptable for second- or first-line therapy.

For example, in the case of the prescription of pembrolizumab in the UK, first-line treatment can be administered with a TPS at least 50%, and second-line with a TPS at least 1%; in both cases, prescription requires that the presence of a sensitising EGFR mutation or ALK translocation be excluded.

PD-L1 expression assessment is now established in routine practice, but is nonetheless not without challenges. It is known that PD-L1 shows heterogeneity in its expression—both spatial and temporal. This has particular implications for NSCLC, where the only histological specimen from a patient may be a small biopsy: sampling of only a negative area of an otherwise positive tumour might deny a patient beneficial therapy. It has also been shown that substantial differences in PD-L1 expression exist between primary and metastatic tumour sites (14), that the method of sampling may have a bearing on the final result (15), and that PD-L1 expression may be influenced by prior chemotherapy (16). As a biomarker, therefore, PD-L1 is very much a “moving target” and a single assessment provides only a snapshot of PD-L1 expression in a single place and a single time. Given that the clinical trials which established the efficacy of checkpoint inhibitors used tissues for PD-L1 testing from a range of tumour sites and acquired by a variety of methods, there exists no strong clinical data to suggest which tissue should be tested for PD-L1 expression; at present, therefore, we are in the unenviable position of not knowing which tissue should be sampled, when it should be sampled, and how it should be sampled, to maximise its potential to predict therapeutic response.

The advent of anti-PD-L1/PD-L1 therapy and of PD-L1 as a biomarker has transformed the landscape of NSCLC management, and has provided an efficacious therapy for patients whose only recourse would otherwise be chemotherapy. Nonetheless, it must be appreciated that PD-L1 expression is far from a perfect biomarker. In the Keynote-024 trial, objective responses were seen only 44.8% of patients whose tumours expressed PD-L1 at least 50% (6); though this is a dramatic improvement on the 27.8% response rate to chemotherapy, it suggests that PD-L1 expression is only part of the story.

Tumour mutation burden assessment

By definition, tumour cells are genetically unstable and harbour high numbers of somatic mutations. Nonetheless, between types of malignant neoplasms there is considerable variation in the prevalence of such mutations, with NSCLC being amongst the most heavily-mutated (17). Owing to a degree of redundancy in transcription and translation, some of these mutations—known as synonymous mutations—will ultimately lead to production of the normal protein; a proportion, however, are non-synonymous and will result in translation of an abnormal protein. As such, tumours with high non-synonymous TMB such as NSCLC express large numbers of abnormal proteins, which are recognised by the immune system as neoantigens (18). It has elegantly been hypothesised that these tumours are potentially amenable to immune-mediated destruction if their immune checkpoints could be circumvented.

Early evidence has suggested that this hypothesis may, indeed, be accurate. Retrospective analyses have shown that response to checkpoint inhibitor therapy in NSCLC is indeed correlated with TMB (19,20). A later study examined the relationship between nivolumab response and both TMB and PD-L1 expression; it demonstrated that high-TMB/high-PD-L1 tumours showed better responses to checkpoint inhibitor therapy than to chemotherapy but—more interestingly—showed improved checkpoint inhibitor therapy responses than low-TMB/any-PD-L1 and high-TMB/low-PD-L1 tumours (21).

Though these data are promising, the major difficulty hitherto has been devising a practical means of assessing TMB. Trawling the genome for each and every mutation would certainly allow TMB assessment, but is both time- and tissue-intensive. The challenge has been to identify a set of genes which is sufficiently large and exhaustive as to act as a proxy for the entire genome, while being sufficiently parsimonious as to be assessable in routine practice.

More recently, the CheckMate-227 trial has reported results based on TMB assessment using the FoundationOne CDx assay, using formalin-fixed, paraffin-embedded tumour samples to assess for mutations in 324 genes. A cut-off of 10
mutations per megabase was used to distinguish low from high TMB. It was shown on this basis that combination nivolumab-ipilimumab checkpoint inhibitor therapy was associated with significantly better response than standard chemotherapy in high-TMB tumours (22). Significantly, this was independent of PD-L1 expression. Although speculative, this raises the possibility that at least some of the high-PD-L1, non-immunotherapy-responsive tumours may have low TMB.

Interestingly, there is evidence that the predictive power of TMB is limited not only to NSCLC; there is increasing evidence to suggest that TMB can meaningfully predict checkpoint inhibitor response across a range of solid malignancies (23,24). In this sense, therefore, it is likely that TMB represents an “agnostic biomarker”, having predictive power regardless of tumour type. To date, however, NSCLC is the only tumour type in which TMB has been clinically validated as having predictive power.

After the fashion of EGFR mutation testing, the possibility of plasma TMB testing using circulating tumour DNA is an attractive one. At present, at least, this is technically not feasible: in a given plasma sample, it is impossible to determine whether detected DNA is derived from tumour or from non-tumour cells and so it is not possible to ascribe a particular detected mutation to the tumour in question.

**MMR deficiency (MMRd)/microsatellite instability (MSI) assessment**

MMR is a molecular system which operates in normal human cells to repair inaccuracies introduced in new DNA strands in the course of DNA replication. In various situations—inherited and sporadic—this system can become dysfunctional and can result in the aggregation of large numbers of mutations. Amongst them, expansions of short, repetitive sequences of DNA called microsatellites are the most common. This state is called MSI, and results in the accumulation of extremely large numbers of somatic mutations. The terms “MMRd” and MSI are therefore interchangeable.

MMRd is best-established in the context of colorectal carcinoma, where it has been shown that the presence of MMRd is strongly associated, firstly, with TMB and, secondly, with favourable responses to pembrolizumab (25). However, the predictive power of MSI is limited not only to colorectal carcinoma. On the basis of a series of Keynote trials demonstrating the relationship between MMRd and pembrolizumab response in a variety of tumour types, the FDA in 2017 granted accelerated approval to the use of the drug in any solid malignancy with MMR defects, progressing following prior treatment and with no other satisfactory treatment options. MSI is therefore the first example of a truly agnostic predictive marker for targeted therapy.

In this sense, MMRd is closely related to TMB: MMRd is a major means in human cancers of developing large numbers of somatic mutations. Given this, it remains uncertain whether the ability to assess TMB renders MMR assessment redundant in the context of immunotherapy. It is certainly true that immunohistochemical assessment of MMR protein expression is extremely well-established, and could easily be put to use as a rapid and affordable test in the context of lung cancer; in this sense, MMR assessment has an edge—at least at present—over TMB assessment.

However, a more important consideration is that MMRd is not the only means of a tumour cell developing high TMB. The frequency of high TMB in NSCLC, for example, is likely a result of the mutagenic effects of exogenous carcinogens in tobacco; indeed, it has been shown that smoking status is associated with response to pembrolizumab (26). Indeed, the demonstration that MMRd occurs in less than 1% of lung adenocarcinomas (27) strongly suggests that sole MMR testing will miss the vast majority of TMB-high NSCLC. On this basis, therefore, the role of MMR assessment in NSCLC is likely to be very limited.

**CD8 T-cells**

In order for the immune system to recognise a malignant cell, it is necessary that the cell has high TMB (and so expresses a sufficient number of neoantigens) and that it expresses checkpoint inhibitors at a low level. However, these two factors are immaterial if there is no immune system presence to drive the anti-tumour immune response. In the case of anti-tumour activity, the key actor in the immune response is the CD8-positive cytotoxic T-cells.

The early evidence for the prognostic and predictive role of CD8 T-cells derives from melanoma, in which there is evidence that high densities of tumour-infiltrating lymphocytes are associated with improved prognosis (28) and that this may also have predictive value in the context of immunotherapy (29).

As in melanoma, there is evidence that CD8 T-cell density is a prognostic marker in NSCLC. There is now
reasonable evidence that increased density of tumour-associated CD8-positive T-cells is associated with improved prognosis in NSCLC (30-33). To date, however, there is no compelling evidence to suggest that increased numbers of tumour-infiltrating CD8 T-cells are associated with improved response to checkpoint inhibitors. A very recently-published study, however, has demonstrated that increased numbers of peripheral blood CD8 T-cells after administration of pembrolizumab is associated with improved treatment response in NSCLC (34). While this certainly does not prove a predictive role for intratumoral T-cells, it certainly makes it plausible.

The extent to which CD8 T-cell infiltration assessment will become routine in clinical practice is unclear; it is certainly fair to say that the underlying evidence has lagged substantially behind that of PD-L1 expression and TMB. In part, this may relate to methodological difficulties in clinically validating the prognostic and predictive implications of CD8 T-cell infiltration. Much like PD-L1, it is probable that such infiltrates demonstrate substantial spatial and temporal heterogeneity, which would make determination of anything other than large prognostic/predictive impacts difficult to establish.

**The future: combinations of markers, combinations of therapies**

In the immunotherapeutic workup of NSCLC, much remains to be proven conclusively. However, what can be stated with confidence is that no single marker is likely to be sufficient to provide all the necessary predictive information for immunotherapy response. PD-L1 expression assessment is firmly established in routine practice, and its predictive power is likely to be refined in future by TMB and—possibly—by CD8 T-cell assessment.

Immunotherapy far from the whole picture, however. Targeted therapies against molecular drivers offer the greatest prospect of long survival in NSCLC, and so future testing algorithms will need to incorporate EGFR mutation, ALK translocation and ROS1 translocation testing. Furthermore, the fact that histological tumour type, stage and patient fitness—above all else—are the most important determinants of the most appropriate treatment cannot be overlooked.

This combination of patient fitness, tumour histology, tumour stage, molecular drivers and immunotherapeutic markers makes for an astonishingly complex management algorithm, which must be informed by clinically-validated trial data. It is clear that—under such circumstances—the question, “how should I treat this patient?”, may well be unanswerable by simple recourse to guidance or common-sense. Indeed, it is probable that such complexity will require computerised algorithmic decision-making, using probabilistic calculations based on trial data (Figure 2).

As if that were not complex enough, it is clear that the sharp divisions between classical chemoradiotherapy, targeted molecular therapy and immunotherapy are beginning to become blurred. Recent data, for example, suggest that combination pembrolizumab-chemotherapy is superior to pembrolizumab monotherapy (35), and ongoing studies are exploring the efficacy of combination anti-EGFR tyrosine kinase inhibitor-checkpoint inhibitor therapy in NSCLC (36).

Clearly, the need for morphological, immunohistochemical, targeted molecular and immunotherapeutic tests places considerable demands on tissue. This would not be an issue if it were the case that all NSCLC specimens underwent resection prior to testing, but this is not at all the case: indeed, many such cancers will never be resected, and the only tissue ever acquired may be a small biopsy. Given that
patients with NSCLC often have significant comorbidity and may not be able to undertake numerous invasive procedures, it is incumbent on the pathologist to ensure that maximal use is made of whatever tissue is available. In part, this is a technical issue which can partially be overcome by the development and use of technologies which use tissue as efficiently as possible. In part, though, it also requires that practitioners be judicious in their use of tissue, keeping wastage to an absolute minimum.

Not only this, but time also poses a great challenge. Patients with NSCLC have a poor prognosis and, if treatment is to be at its most effective, it must be instituted as rapidly as possible. Furthermore, because the various markers of interest are interrelated from management points of view, all results have to be available rapidly; there is at present, for example, little point in providing a very timely PD-L1 expression result if an EGFR mutation result takes weeks, because according to licencing anti-PD-1/ PD-L1 therapy requires the absence of a sensitising EGFR mutation. Again, of course, improvements in technologies can help, but above all else a robust logistical setup is essential.

Of great interest in the last few years has been the prospect of testing circulating tumour DNA in plasma for molecular alterations. This is established in the case of EGFR mutations in NSCLC as a surrogate or complement to tissue testing at the time of diagnosis, or clinical progression under tyrosine kinase inhibitor therapy to detect a secondary T790M mutation. It remains to be demonstrated that such a test can be used to monitor patients under treatment. With regards wider mutation screening which could be used to assess TMB, no data currently exist to demonstrate its clinical utility; the major challenge in this case is the small amount of circulating tumour DNA, the lack of technique to distinguish tumour-derived from non-tumour-derived DNA, and the level of sensitivity required to be clinically relevant. Prospective trials, perhaps, hold the key in clinically validating gene mutation testing on plasma.

**Conclusions**

As described, much remains uncertain in the field of immunotherapy in NSCLC. What can be assured, however, is that one single test will not provide all the information required for effective patient management. Rather, reflecting the clinical complexity combination multi-modality therapies (e.g., immuno-chemotherapy), the emergence of complex algorithms incorporating information derived from multiple tests performed simultaneously will achieve paramount importance in coming decades.

**Acknowledgements**

None.

**Footnote**

**Conflicts of Interest:** The authors have no conflicts of interest to declare.

**References**


