



OPEN ACCESS

Improving practice in PD-L1 testing of non-small cell lung cancer in the UK: current problems and potential solutions

John R Gosney,¹ Michael D Peake,^{2,3} Keith M Kerr⁴

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/jcp-2022-208643>).

¹Cellular Pathology, Royal Liverpool and Broadgreen Hospitals NHS Trust, Liverpool, UK

²Center for Cancer Outcomes, North Central and North East London Cancer Alliances, UCLH, London, UK

³Grobby Road Hospital, University of Leicester, Leicester, UK

⁴Pathology, University of Aberdeen, Aberdeen, UK

Correspondence to

Professor Keith M Kerr, Pathology, University of Aberdeen, Aberdeen AB25 2ZD, UK; k.kerr@abdn.ac.uk

Presented, in part, as a poster presentation (abstract number #85) at the British Thoracic Oncology Group (BTOG) 19th Annual Conference held online on 22nd and 23rd April 2021.

Received 7 November 2022

Accepted 20 December 2022

Published Online First

5 January 2023



© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Gosney JR, Peake MD, Kerr KM. *J Clin Pathol* 2024;**77**:135–139.

ABSTRACT

Aims Programmed cell death ligand 1 (PD-L1) expression, used universally to predict response of non-small cell lung cancer (NSCLC) to immune-modulating drugs, is a fragile biomarker due to biological heterogeneity and challenges in interpretation. The aim of this study was to assess current PD-L1 testing practices in the UK, which may help to define strategies to improve its reliability and consistency.

Methods A questionnaire covering NSCLC PD-L1 testing practice was devised and members of the Association of Pulmonary Pathologists were invited to complete this online.

Results Of 44 pathologists identified as involved in PD-L1 testing, 32 (73%) responded. There was good consistency in practice and approach, but there was wide variability in the distribution of PD-L1 scoring. Although the proportions of scores falling into the three groups (negative, low and high) defined by the 1% and 50% 'cut-offs' (38%, 33% and 27%, respectively) reflect the general experience, the range within each group was wide at 23–70%, 10–60% and 15–36%, respectively.

Conclusions There is inconsistency in the crucial endpoint of PD-L1 testing of NSCLC, the expression score that guides management. Addressing this requires formal networking of individuals and laboratories to devise a strategy for its reduction.

INTRODUCTION

Lung cancer remains the leading cause of cancer-related deaths in the UK in both men and women.¹ This is despite the fact that the majority of these tumours, those classified as non-small cell lung cancer (NSCLC), comprise the group for which an increasing range of targeted therapies has been developed over the past decade.² Such therapies include an expanding group of tyrosine kinase inhibitors targeted against tumours with specific genetic aberrations, that is, single genomic drivers, and a group of immune-modulating drugs (IMs) targeted against the programmed cell death protein 1 (PD-1)/programmed cell death ligand 1 (PD-L1) immune checkpoint.^{3,4}

The IMs depend for their efficacy on the tumour exploiting the PD-1/PD-L1 checkpoint to protect itself from an immune response, an adaptive mechanism that manifests itself in increased expression of PD-L1 on the surface of tumour cells.⁵ A range of IMs is currently approved in the UK for the treatment of NSCLC (table 1), differing in terms of their licensed indication, as defined by

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Advances in first-line and second-line therapy have led to the approval of immune-modulating drugs for patients with non-small cell lung cancer (NSCLC). Programmed cell death ligand 1 (PD-L1) expression offers a predictor of response for many of these medicines, but it is a fragile biomarker and there is a pressing need for greater consistency in its reporting across laboratories.

WHAT THIS STUDY ADDS

⇒ Assessment of current PD-L1 testing practice in the UK provides new understanding of the variability observed between centres, particularly in the distribution of PD-L1 scoring.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The survey results evidence the need for formal networking of individuals and laboratories to reduce inconsistency in the assessment and reporting of the expression score, the crucial endpoint of PD-L1 testing in NSCLC.

the European Medicines Agency, and in patient eligibility, as defined by the National Institute for Health and Care Excellence and Scottish Medicines Consortium. Among these eligibility criteria is the level of expression of PD-L1, as detected by immunohistochemistry (IHC). This is generally reported as the tumour proportion score (TPS), the percentage of tumour cells expressing PD-L1 on their surface.⁵

Assessing expression of PD-L1 is by far the most commonly used predictor of response of NSCLC to IMs. Unfortunately, it is a fragile biomarker, compromised by its biological heterogeneity, variations in laboratory practice, including reluctance to use 'cytology' specimens for its assessment, and challenges in interpretation.^{6–11}

Many of these challenges can be addressed only by understanding the nature of and variability in the practice and experience of those involved in PD-L1 testing across a wide range of laboratories and this information is not currently available on the necessary scale. As a precursor to defining a strategy to improve the reliability and consistency of PD-L1 testing, we thought it essential to gather comprehensive data on current practice across the UK.

Table 1 PD-1/PD-L1 blockade therapy in advanced NSCLC: options available in the UK as of October 2022

Manufacturer	INN (brand name)	Target and IgG isotype	Approved treatment for NSCLC by EMA	UK patient eligibility (NICE TA guidance)
Merck Sharp & Dohme	Pembrolizumab (KEYTRUDA)	Anti-PD-1 (humanised IgG4/kappa)	Monotherapy (1L, stage IV)	PD-L1 TPS $\geq 50\%$ NSCLC that has no <i>EGFR</i> + or <i>ALK</i> + mutations (NICE TA531)
			Pembrolizumab + chemotherapy (1L, stage IV)	Regardless of PD-L1 status in non-squamous NSCLC that has no <i>EGFR</i> + or <i>ALK</i> + mutations (NICE TA683); and regardless of PD-L1 status in squamous NSCLC (NICE TA770)
			Monotherapy (2L, stage III/IV)	PD-L1 TPS $\geq 1\%$ (NICE TA428)
Bristol-Myers Squibb	Nivolumab (OPDIVO)	Anti-PD-1 (human IgG4)	Nivolumab + ipilimumab + chemotherapy (1L, stage IV)	Not recommended for NSCLC that has no <i>EGFR</i> + or <i>ALK</i> + mutations (NICE TA724)
			Monotherapy (2L, stage III/IV)	In squamous NSCLC: regardless of PD-L1 status (NICE TA655); in non-squamous NSCLC: only if PD-L1 TPS is $\geq 1\%$ (NICE TA713)
Roche/ Genentech	Atezolizumab (TECENTRIQ)	Anti-PD-L1 (humanised IgG1)	Monotherapy (1L, stage IV)	NSCLC having PD-L1 TPS $\geq 50\%$ or $\geq 10\%$ of tumour-infiltrating cells and that has no <i>EGFR</i> + or <i>ALK</i> + mutations (TA705)
			Atezolizumab + bevacizumab + chemotherapy (1L, stage IV)	In non-squamous NSCLC only: PD-L1 TPS between 0% and 49% (NICE TA584)
			Atezolizumab + chemotherapy (1L, stage IV)	No recommendation made for non-squamous NSCLC (NICE TA618)
			Atezolizumab + bevacizumab + chemotherapy (2L, stage IV)	After failure of targeted therapy for <i>EGFR</i> + or <i>ALK</i> + non-squamous NSCLC, regardless of PD-L1 status (NICE TA584)
AstraZeneca	Durvalumab (IMFINZI)	Anti-PD-L1 (human IgG1 kappa)	Monotherapy (2L, stage III/IV)	Regardless of PD-L1 status (NICE TA520)
			Monotherapy (after completion of chemoradiation, stage III)	PD-L1 TPS $\geq 1\%$ and only when disease has not progressed after platinum-based chemoradiation (NICE TA798)

ALK, anaplastic lymphoma kinase; EGFR, epidermal growth factor receptor; EMA, European Medicines Agency; INN, international non-proprietary name; 1L, first line; 2L, second line; NICE, National Institute for Health and Care Excellence; NSCLC, non-small cell lung cancer; PD-L1, programmed cell death ligand; TA, technology appraisal; TPS, tumour proportion score.

MATERIALS AND METHODS

A questionnaire was devised covering many aspects of PD-L1 testing of NSCLC, and members of the Association of Pulmonary Pathologists (APP; appathologists.com) were contacted by email and invited to participate. The APP has a broad membership, ranging from general pathologists in district general hospitals who have an interest in the area to single-speciality pathologists in academic institutions, many of whom service tertiary thoracic surgical centres. To ensure the capture of as many testing centres as possible, the APP membership list was checked against other contact lists of laboratories and individual pathologists known by the authors to be involved in PD-L1 testing of NSCLC. To avoid duplication, participants were requested to complete the survey only if they were the lead person at their centre responsible for testing. All responses were anonymous to encourage participation and open disclosure.

The survey comprised 26 questions. For the majority of these (19 of the 26), respondents selected a response from prespecified options. These covered such areas as the sources, number and nature of specimens tested, by whom this was performed and their involvement in thoracic pathology in general and in PD-L1 testing specifically, at what point in the diagnostic and management pathway testing was performed, the assay(s) used, turnaround times (TATs) between receipt of samples in the laboratory and reporting of results, and training and involvement in external quality assurance (EQA) schemes. In four of these areas, a more nuanced free-text response was requested: reflex testing, expression and reporting of results, approach to repeating a test, and the range of results obtained across three groups as determined by PD-L1 expression scores ('negative', 'low' and 'high') according to the conventional 1% and 50% 'cut-offs'. The survey remained open between 12 June 2020 and 17 July 2020.

RESULTS

Of the 44 centres approached, a pathologist primarily involved in PD-L1 testing of NSCLC responded from 32 (72.7%); 25 of these respondents (78.1%) responded to questions up to the final question (although some of these respondents did not reply to all of the 26 questions). The responses to questions requiring only selection of a prespecified response are shown online (online supplemental file 1).

The free-text responses can be summarised as follows:

For reflex testing, centres receiving specimens from a variety of sources had no control of how the decision to test was being made, but the details of the process varied widely. Occasional perceptions acting against reflex testing included that many patients are unsuitable for IM therapy anyway on the grounds of performance status, and that securing reimbursement for it might be problematic.

The approach to expression and reporting of results showed some variation across centres: 48% expressing them as the TPS and 37% as within a 'categorical range' (ie, $<1\%$, 1%–49% or $\geq 50\%$). One reported them as $<1\%$, 1%–5% and then at 10% intervals 'as agreed with oncologists'. No centre described the result merely as 'low/high' or 'negative/positive'.

The approach to retesting was largely consistent across centres. All would test a second specimen, if available, when a previous specimen had been inadequate (<100 tumour cells).⁵ A second specimen was often tested even in the context of a previous successful assessment, either on disease progression or because the result of the initial test had been very close to one of the crucial 'cut-off' points. Occasionally, an oncologist would request testing of a further specimen from the same tumour site if the initial test on an adequate specimen had been 'negative', but they were 'running out of options'—the inference being that a second test might yield a higher score.

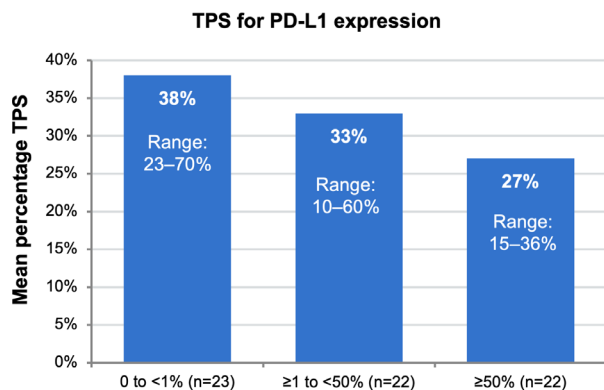


Figure 1 TPS for PD-L1 expression in NSCLC samples at respondents' centres. Mean percentage of NSCLC samples that have a TPS of 0 to <1%, ≥1 to <50%, or ≥50% for PD-L1 expression. NSCLC, non-small cell lung cancer; PD-L1, programmed cell death ligand 1; TPS, tumour proportion score.

The range of results obtained showed unexpected variation across centres (figure 1). Within each of the three categories defined by the usual 'cut-off' points, and which are often referred to as 'negative' (0% to <1%), 'low' (1%–49%) and 'high' (≥50%), variation was wide at 23%–70%, 10%–60% and 15%–36%, respectively.

DISCUSSION

Assessment of PD-L1 expression as detected by IHC is currently the only 'test' used universally to guide the prescription of IMs to treat patients with NSCLC, and its implementation has not been straightforward. Variation in specimen processing and in the experience of pathologists engaged in its interpretation augment the unavoidable challenges inherent in its biology and weaken its predictive power. The results of our survey highlight well this variability and raise the obvious question of how it might be reduced. In the context of UK practice, we believe our survey to be the most comprehensive yet performed in this area of diagnostics, in terms of coverage of those active in this area and the data collected. A more detailed understanding of why such variability exists is a prerequisite to devising a strategy to reduce it, assuming that variability is detrimental to the desired endpoints.

Laboratory practice

Variability in laboratory practice (the handling, processing and preparation of specimens preassessment) is almost a tradition in pathology, a legacy of an approach that, until recently, owed more to cookery than to uniform, evidence based, regulated and tightly controlled practice. Such variability was highlighted in a recent review addressing the use of cytology specimens for assessing PD-L1 expression in NSCLC¹¹ and is important because its ultimate consequence is that specimens prepared by different laboratories might already vary in how PD-L1 expression is manifested before they are interpreted by a pathologist. Such variability has been brought into sharp focus by the increasing requirement for broader predictive 'biomarker testing' of NSCLC using IHC, and by studies showing how variation in such techniques can have an impact on treatment choices. This is illustrated, for example, by the results reported by the UK National External Quality Assessment Service (NEQAS) on assessing expression of anaplastic lymphoma kinase (ALK) fusion protein.¹² The ready availability of EQA schemes, across

the developed world at least, provides an obvious mechanism for standardising laboratory practice and reducing variability.¹³ A comparison can be drawn between the current situation with PD-L1 testing and the serious variability in the technical quality of specimens of breast cancer assessed for human epidermal growth factor receptor 2 (Her2) expression that became apparent in the early 2000s when UK NEQAS established an EQA scheme specifically for this predictive test.^{14 15} A similar scheme for PD-L1 expression in NSCLC is now well established by UK NEQAS and is generating valuable information about interlaboratory variability; in the UK, subscription to such schemes is mandatory for laboratories performing such analyses in order for them to obtain UK Accreditation Service accreditation (standard ISO15189).⁶ It is important, however, that this information is acted on and the effect of these improvements re-audited. It is sobering also to realise that, in many countries, subscription to such EQA schemes is not mandatory.

Interpretation

Identifying the reasons for, and then improving interpretation of, PD-L1 expression by pathologists is more challenging still. The most worrying result of our survey is the wide variability of scoring PD-L1 expression within the three broad groups, 'negative' (0%–1%), 'low' (1%–49%) and 'high' (≥50%). These scores, the ultimate endpoints of PD-L1 testing on which crucial clinical decisions are made, should show relatively limited variation between centres since it is unlikely, in the context of UK patients with NSCLC, that significant variation in the range of PD-L1 expression will occur for reasons of biology or geography. It is well established from clinical trials and other reports that the distribution of PD-L1 TPSs is approximately even across the three categories of 'negative', 'low' and 'high' with, perhaps, a tendency for slightly fewer cases in the middle category, leading towards a bimodal distribution.^{5 7–11} Broadly speaking, therefore, there is evidence from our survey that some centres may be 'under-reporting' the PD-L1 TPS. With the deployment of stage-agnostic reflex testing, which appears to be the dominant approach in this survey of UK centres, there could be a slight bias towards a greater, though still relatively small, proportion of early-stage disease in the test population when compared with data from clinical trials of patients with more advanced disease. Although there is evidence for lower PD-L1 expression in early stage disease,¹⁶ this still would not account for the 'outliers' in this survey reporting high proportions of specimens as 'negative'. Most of the laboratories in our survey used trial-validated companion diagnostic assays, so it is unlikely that the observed variation is due to poor assay sensitivity.

Of course, there will always be some variability; interpreting PD-L1 expression is, by its very nature, subjective, but we do not believe that the variability we reveal here is acceptable. Guidelines for which pathologists should and should not interpret PD-L1 expression in NSCLC have emerged over recent years, but are difficult, if not impossible, to enforce. It has been suggested, for example, that interpretation should be restricted to pathologists who see at least 200 diagnostic lung cancer specimens a year, have undergone appropriate formal training (which results in some evidence of competence) and subscribe to an appropriate EQA scheme that is interpretative, not technical.⁷ Even among the laboratories covered by our survey, in which at least one pathologist, as a member of the APP, clearly has an interest in thoracic pathology, there are some worrying trends. For example, more than a third of laboratories handle fewer than five PD-L1 tests a week and, in more than 15%, the PD-L1

testing workload is spread between five and eight pathologists (online supplemental file 1).

All pathologists involved in PD-L1 scoring are aware of how difficult it can be and of its subjectivity. In the training programmes that are delivered for PD-L1 assessment by means of a TPS, emphasis is put on how to (semi)quantify, if not actually count, the number of tumour cells in the sample and the proportion that are 'positive'. All levels of staining intensity are relevant and are counted. In a proportion of cases, staining can be weak, requiring examination at high magnification. As pathologists become more familiar with an assay such as PD-L1 scoring, the time required for each assessment will inevitably reduce. Anecdotally, we also hear reports of a more 'gestalt' approach to assessment that could conceivably lead to small numbers of positive cells, or cells with light staining, being missed. As many pathologists are currently practising under pressurised conditions with poor staff/workload ratios and pressure to improve TAT, taking such shortcuts is understandable; more than a quarter of respondents in this survey reported average TATs of 5 days or more.

In comparison with clinical trials, from which cytology specimens were excluded, it is difficult to know precisely what impact the regular, routine testing of such specimens might have had on our observed outcomes. Most pathologists acknowledge that, in general, PD-L1 scoring of cytology specimens can be challenging and require more time, but there is no conclusive evidence that PD-L1 scores per se are lower in cytology as compared with histology ('biopsy') specimens.^{7 10 11} As discussed above, there is considerable variability in how cytology specimens are processed, and this may well contribute to variability in the results obtained from their assessment.¹⁷

In view of these challenges, there is growing interest, as in other difficult areas of diagnostic pathology, in the use of image analysis, algorithms and machine learning as an aid to interpretation. For example, the validation of such software as an aid to interpretation of PD-L1 expression in NSCLC is a component of the Northern Pathology Imaging Co-operative project,¹⁸ which is currently assessing its utility to a range of pathologists with varying levels of experience across six universities in the North of England.

Some variability is inevitable in such complex systems as laboratories in which activity is run and undertaken by individuals who vary in their approach, practice and the variety of skills they possess, and is not surprising. Indeed, a very similar pattern of variability, although in a slightly different context, was revealed by the LungPath study.¹⁹ In this survey, the approach of laboratories and pathologists to subclassifying NSCLCs into squamous and adenocarcinoma was examined, and the findings are largely recapitulated by those we describe here. This is not to say, however, that such variability cannot be reduced.

We suggest that a formal network is established of all laboratories engaged in PD-L1 testing of NSCLC with a view to sharing details of practice and data resulting from testing. This would provide a basis for standardising and improving practice and would carry an important educational component.

Ultimately, however, encouraging and supporting adoption of best practice might require a more rigorous approach by those institutions, such as the Royal College of Pathologists and Institute of Medical Laboratory Scientists, that are responsible for training, examining and maintaining standards. Part of the approach to remedying the serious inconsistencies in assessing specimens of breast cancer for Her2 expression referred to above consisted of removing the service from 'failing' laboratories. This

greatly improved quality and consistency and set an important precedent.

Adequacy of samples

The only objective metric we have for sample adequacy for PD-L1 testing is the presence of at least 100 viable tumour cells in the tissue section being assessed. Intuitively, this makes sense when one is delivering a percentage score on a sample that is already severely challenged by biological heterogeneity and sampling 'error' but raises questions about how representative of the patient's disease burden the rendered score actually is. There is evidence that TPSs reported on samples that have <100 tumour cells are much less predictive of response to IMs than scores derived from samples that are richly cellular.²⁰ It is comforting that awareness and reporting of this criterion of sufficiency seems to be universal in our survey.

Our survey is by no means the first to highlight the problems and challenges with PD-L1 testing in NSCLC, which were clearly apparent, for example, in the global survey conducted by the Pathology Committee of the International Association for the Study of Lung Cancer.²¹ However, we wished to concentrate specifically on practice in the UK so that addressing and resolving any problems that might become apparent could be managed efficiently under the auspices of the APP, which is a UK-based association with strong national links.

It is gratifying, for example, that the College of American Pathologists is currently in the process of developing guidelines for PD-L1 testing of patients with lung cancer in an attempt to standardise and improve assessment, a strategy that also considers the possible utility of assessing tumour mutational burden as an adjunctive investigation.²²

It is always politically difficult to impose what are often interpreted as restrictions on what individuals might or might not do, even to the point of their being seen as a threat to individuality. In the end, however, the only significant measure of quality of any test we perform, or assessment we make, is arriving at the right answer for the patient, the ultimate user of the service we provide.

CONCLUSION

There is clearly inconsistency in the assessment and reporting of the expression score, the crucial endpoint of PD-L1 testing in NSCLC, that is central to guiding patient management. Addressing this requires formal networking of individuals and laboratories to devise a strategy for reducing this variation.

Handling editor Runjan Chetty.

Acknowledgements The authors gratefully acknowledge insights and feedback from the following experts who (along with the authors) participated in an online advisory board meeting in September 2020, which focused on PD-L1 testing in pulmonary pathology in the UK: Dr Thomas Newsom-Davis, Chelsea and Westminster Hospital; Dr Shobhit Bajjal, University Hospitals Birmingham NHS Foundation Trust; Dr Matthew Evison, Wythenshawe Hospital; and Mr Ehab Bishay, Heart of England NHS Foundation Trust. Medical writing and editorial assistance in the preparation of this manuscript, which was in accordance with Good Publications Practice (GPP) guidelines, was provided by Patrick Foley, PhD, of NexGen Healthcare Communications (London, UK) and funded by AstraZeneca UK.

Contributors JRG and KMK conceptualised the study, designed the protocol and questionnaire, and analysed the data. JRG wrote the original draft manuscript. All authors edited and revised the manuscript and controlled the decision to publish. JRG acts as guarantor for this work and the conduct of the study.

Funding The UK PD-L1 online survey described in this manuscript, the subsequent advisory board meeting, presentation of the survey results at the BTOG 2021 meeting and medical writing assistance for the development of this manuscript were supported by AstraZeneca UK.

Competing interests JRG is a paid speaker for, advisor to, or has received research support from AbbVie, Amgen, AstraZeneca, Bayer, Boehringer Ingelheim, Bristol-Myers Squibb, Diaceutics, Guidepoint, Lilly & Co, Merck Sharp & Dohme, Novartis, OncLive, Pfizer, Roche, and Takeda Oncology. MDP has received honoraria for participation in advisory boards from AstraZeneca. KMK has received consultancy and/or speaker's fees from AbbVie, Amgen, Archer Diagnostics, AstraZeneca, Bayer, Boehringer Ingelheim, Bristol-Myers Squibb, Debiopharm, Diaceutics, Eli Lilly, Medscape, Merck Serono, Merck Sharp & Dohme, Novartis, PeerVoice, Pfizer, Prime Oncology, Regeneron, Roche, Roche Diagnostics/Ventana.

Patient consent for publication Not applicable.

Ethics approval This study did not require ethics committee approval.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

- 1 Cancer Research UK. Lung cancer statistics, 2020. Available: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/lung-cancer#heading-One> [Accessed 6 Oct 2022].
- 2 Yuan M, Huang L-L, Chen J-H, et al. The emerging treatment landscape of targeted therapy in non-small-cell lung cancer. *Signal Transduct Target Ther* 2019;4:61.
- 3 Liu T, Ding S, Dang J, et al. First-line immune checkpoint inhibitors for advanced non-small cell lung cancer with wild-type epidermal growth factor receptor (EGFR) or anaplastic lymphoma kinase (ALK): a systematic review and network meta-analysis. *J Thorac Dis* 2019;11:2899–912.
- 4 Pennell NA, Arcila ME, Gandara DR, et al. Biomarker testing for patients with advanced non-small cell lung cancer: real-world issues and tough choices. *Am Soc Clin Oncol Educ Book* 2019;39:531–42.
- 5 Tsao MS, Kerr KM, Dacic S, et al. *IASLC atlas of PD-L1 immunohistochemistry testing in lung cancer*. First. Aurora, CO, USA: Editorial Rx Press, 2017. <https://www.iaslc.org/research-education/publications-resources-guidelines/iaslc-atlas-pd-l1-testing-lung-cancer>
- 6 Cree IA, Booton R, Cane P, et al. PD-L1 testing for lung cancer in the UK: recognizing the challenges for implementation. *Histopathology* 2016;69:177–86.
- 7 Büttner R, Gosney JR, Skov BG, et al. Programmed Death-Ligand 1 immunohistochemistry testing: a review of analytical assays and clinical implementation in non-small-cell lung cancer. *J Clin Oncol* 2017;35:3867–76.
- 8 Evans M, O'Sullivan B, Smith M, et al. Predictive markers for anti-PD-1/PD-L1 therapy in non-small cell lung cancer—where are we? *Transl Lung Cancer Res* 2018;7:682–90.
- 9 Haragan A, Field JK, Davies MPA, et al. Heterogeneity of PD-L1 expression in non-small cell lung cancer: implications for specimen sampling in predicting treatment response. *Lung Cancer* 2019;134:79–84.
- 10 Lantuejoul S, Sound-Tsao M, Cooper WA, et al. PD-L1 testing for lung cancer in 2019: perspective from the IASLC pathology Committee. *J Thorac Oncol* 2020;15:499–519.
- 11 Gosney JR, Boothman A-M, Ratcliffe M, et al. Cytology for PD-L1 testing: a systematic review. *Lung Cancer* 2020;141:101–6.
- 12 Ibrahim M, Parry S, Wilkinson D, et al. Alk immunohistochemistry in NSCLC: discordant staining can impact patient treatment regimen. *J Thorac Oncol* 2016;11:2241–7.
- 13 Dufrainig K, Fenizia F, Torlakovic E, et al. Biomarker testing in oncology - Requirements for organizing external quality assessment programs to improve the performance of laboratory testing: revision of an expert opinion paper on behalf of IQNPath ABSL. *Virchows Arch* 2021;478:553–65.
- 14 Bartlett JMS, Ibrahim M, Jasani B, et al. External quality assurance of HER2 fluorescence in situ hybridisation testing: results of a UK NEQAS pilot scheme. *J Clin Pathol* 2007;60:816–9.
- 15 Dodson A, Parry S, Ibrahim M, et al. Breast cancer biomarkers in clinical testing: analysis of a UK national external quality assessment scheme for immunocytochemistry and in situ hybridisation database containing results from 199 300 patients. *J Pathol Clin Res* 2018;4:262–73.
- 16 Wang H, Agulnik J, Kasymjanova G, et al. Cytology cell blocks are suitable for immunohistochemical testing for PD-L1 in lung cancer. *Ann Oncol* 2018;29:1417–22.
- 17 Bubendorf L, Conde E, Cappuzzo F, et al. A noninterventonal, multinational study to assess PD-L1 expression in cytological and histological lung cancer specimens. *Cancer Cytopathol* 2020;128:928–38.
- 18 Richman SD, Williams CJM, West NP, et al. The N6 Consortium: the Northern pathology network. poster Abstract p21. Molecular and Digital Pathology: The Future, 3rd Joint Meeting with the Royal Society of Medicine 212th Scientific Meeting of the Pathological Society of Great Britain & Ireland, 2020.
- 19 Cane P, Linklater KM, Nicholson AG, et al. Morphological and genetic classification of lung cancer: variation in practice and implications for tailored treatment. *Histopathology* 2015;67:216–24.
- 20 Naito T, Udagawa H, Sato J, et al. A minimum of 100 tumor cells in a single biopsy sample is required to assess Programmed Cell Death Ligand 1 expression in predicting patient response to nivolumab treatment in nonsquamous non-small cell lung carcinoma. *J Thorac Oncol* 2019;14:1818–27.
- 21 Mino-Kenudson M, Le Stang N, Daigneault JB, et al. The International Association for the Study of Lung Cancer global survey on Programmed Death-Ligand 1 testing for NSCLC. *J Thorac Oncol* 2021;16:686–96.
- 22 College of American Pathologists. PD-L1 testing of patients with lung cancer for immuno-oncology therapies, 2021. Available: <https://www.cap.org/protocols-and-guidelines/upcoming-cap-guidelines/pd-l1-testing-of-patients-with-lung-cancer-for-immuno-oncology-therapies> [Accessed 6 Oct 2022].