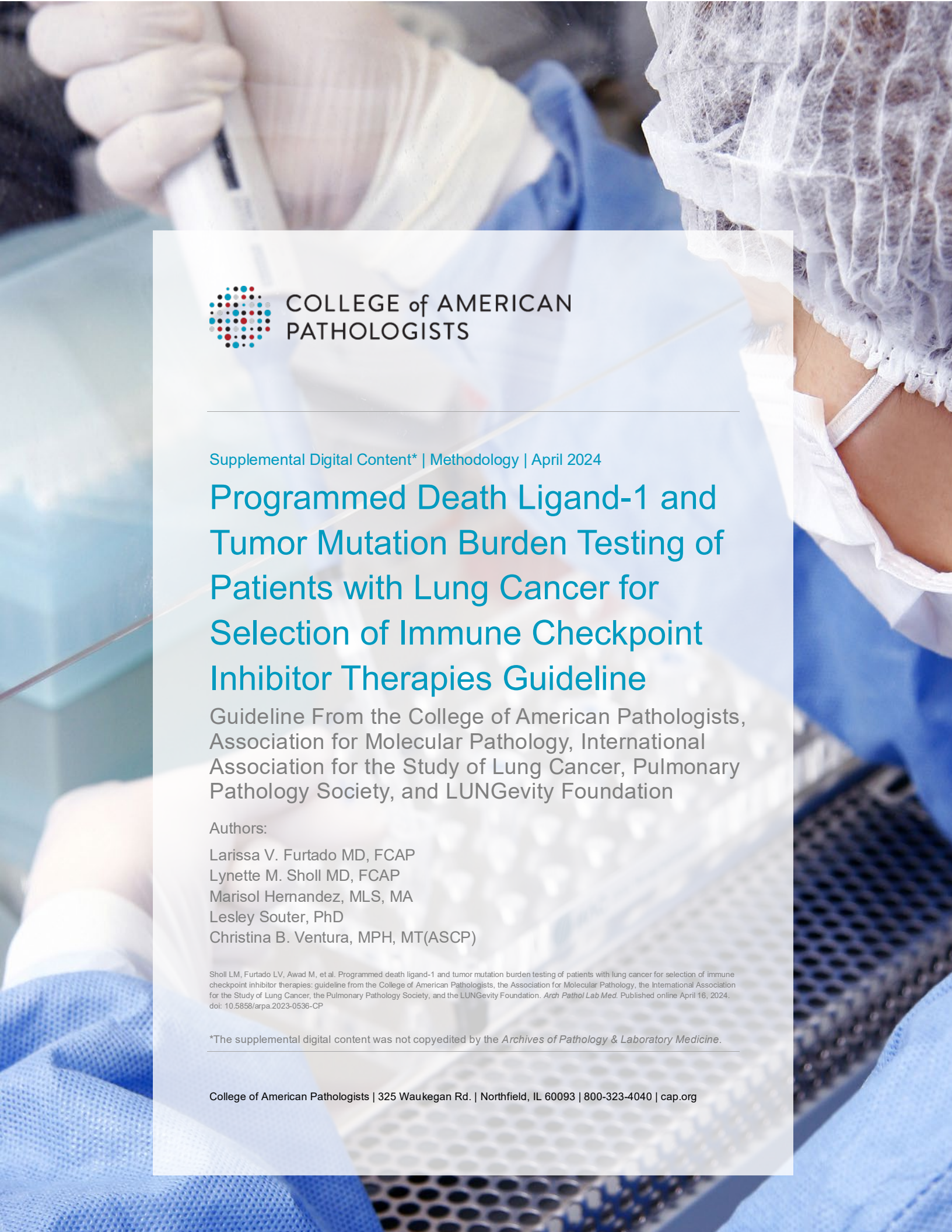


COLLEGE of AMERICAN PATHOLOGISTS

Supplemental Digital Content* | Methodology | April 2024

Programmed Death Ligand-1 and Tumor Mutation Burden Testing of Patients with Lung Cancer for Selection of Immune Checkpoint Inhibitor Therapies Guideline

Guideline From the College of American Pathologists, Association for Molecular Pathology, International Association for the Study of Lung Cancer, Pulmonary Pathology Society, and LUNGevity Foundation

Authors:

Larissa V. Furtado MD, FCAP
Lynette M. Sholl MD, FCAP
Marisol Hernandez, MLS, MA
Lesley Souter, PhD
Christina B. Ventura, MPH, MT(ASCP)

Sholl LM, Furtado LV, Awad M, et al. Programmed death ligand-1 and tumor mutation burden testing of patients with lung cancer for selection of immune checkpoint inhibitor therapies: guideline from the College of American Pathologists, the Association for Molecular Pathology, the International Association for the Study of Lung Cancer, the Pulmonary Pathology Society, and the LUNGevity Foundation. *Arch Pathol Lab Med.* Published online April 16, 2024. doi: 10.5858/arpa.2023-0536-CP

*The supplemental digital content was not copyedited by the *Archives of Pathology & Laboratory Medicine*.

GUIDELINE DEVELOPMENT METHODS

Panel Composition

The College of American Pathologists (CAP) along with its collaborators, the American Society of Clinical Oncology (ASCO), Association for Molecular Pathology (AMP), International Association for the Study of Lung Cancer (IASLC), Pulmonary Pathology Society (PPS), and LUNGevity Foundation convened an expert and advisory panel (EP/AP) consisting of members with experience and expertise in the testing, diagnosis and treatment of patients with lung cancer for selection of immune checkpoint inhibitor (ICI) therapies to develop evidence-based recommendations for programmed death receptor-1 (PD-1) and programmed death ligand-1 (PD-L1) testing. Members include practicing pathologists, clinicians, oncologists, guideline methodologist, and patient advocates from the from the United States and Europe. The CAP approved the appointment of the project cochairs and expert panel members. The following organizations provided official panel representation: ASCO, AMP, IASLC, PPS, and LUNGevity Foundation.

The roles of each panel are described in the Evidence-based Guideline Development Methodology Manual (Methodology Manual).[1]

Conflict of Interest Policy

Prior to acceptance on the expert or advisory panel, potential members completed the CAP conflict of interest (COI) disclosure process, whose policy and form require disclosure of material financial interest in, or potential for benefit of significant value from, the guideline's development or its recommendations 24 months prior through the time of publication. The potential members completed the COI disclosure form, listing any relationship that could be interpreted as constituting an actual, potential, or apparent conflict. A complete description of the COI policy is available in the online Methodology Manual.

Members were required to disclose conflicts prior to beginning and continuously throughout the project's timeline. EP members' disclosures are listed in the appendix of the manuscript. The CAP provided funding for the administration of the project; no industry funds were used in the development of the guideline. All panel members volunteered their time and were not compensated for their involvement, except for the contracted methodologist.

Systematic Evidence Review

The objective of the Systematic Evidence Review (SER) was to identify articles that provided data to inform the recommended testing for the PD-L1 testing of patients with lung cancer for selection of ICI therapies. If of sufficient quality, findings from this review would provide an evidence-base to support the recommendations of the guideline. The scope of the SER and the key questions (KQs) with the PICO elements (Population, Intervention, Comparator, Outcome(s)) were established by the EP in consultation with the methodologist prior to beginning the literature search.

Detailed key questions including the PICO is included in Supplemental Table 1.

Search and Selection

Controlled vocabulary and keyword terms were included to address the key questions. Detailed database search strings are included as Supplemental Figure 1. All search results were deduplicated using reference management software following published methods.[2] All search strategies were reviewed by a second medical librarian using the Press Review of Electronic Search Strategies (PRESS) statement for systematic reviews. Additional searches to supplement the database searches were completed to locate guidelines and unindexed (grey) literature using the following websites: Guidelines International Network, ECRI Guidelines Trust, Trip Medical Database, University of York Centre for Reviews and Dissemination, and relevant US and international organizational websites using the Canadian Agency for Drugs and Technologies in Health (CADTH) Grey Matters document and known pathology organizations' websites.

Selection at all levels was based on the predetermined inclusion/exclusion criteria which are detailed in the manuscript.

Data Extraction & Management

The data elements from an included article/document were extracted by one reviewer into standard data formats and tables developed using the systematic review database software, DistillerSR (Evidence Partners Inc., Ottawa, Canada); a second reviewer confirmed accuracy and completeness. Any discrepancies in data extraction were resolved by discussion between the co-chairs and the methodologist. A bibliographic database was established in EndNote (Thomson Reuters, Carlsbad, CA) to track all literature identified and reviewed during the study.

Outcomes of Interest

According to the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach, it is important for clinical guideline panels to review a comprehensive list of outcomes.[3] The EP was polled to collect information on which outcomes should be included in the PICO. These outcomes included, but were not limited to, accuracy in diagnosis (specificity, sensitivity, positive and negative predictive values), change in patient management, cost, optimal and adequacy of specimen selection, patient preference, quality of life, rates of adverse reactions, survival rates, test/assay utility, and timely communication to the clinicians.

In consideration of the limited scope and resources, the EP ranked the outcomes used in the PICO. Using the GRADE approach[3] of considering the relative importance of outcomes, the EP was polled to rate each initially identified outcome in terms of importance for decision making. The EP voted on a scale of 1–9: outcomes rated 1-3 were defined as “of limited importance”; outcomes rated 4-6 as “important, but not critical”; and outcomes rated 7-9 were “critical for decision making”. The EP finalized the outcomes after a discussion during the first in-person meeting.

Outcomes of Limited Importance*

Note: These outcomes not used for decision making

- Sample adequacy

(*Not used for decision making)

Important Outcomes

- Tissue utilization

Critical Outcomes

- Survival rates (overall survival [OS], disease free survival [DFS], progression free survival [PFS])
- Treatment response rates (objective response rate [RR], complete RR, pathologic RR)
- Adverse events
- Diagnostic test characteristics (sensitivity, specificity, positive predictive value, negative predictive value)
- Clinical and analytical validity (assay concordance, sample concordance, proportion score [tumor proportion score, immune proportion score, combined proportion score])

Strength of Recommendations and Evidence-to-Decision Framework

Development of recommendations required the panel to review the identified evidence and make a series of key judgments using the GRADE approach (Supplemental Table 2). In addition to the panel discussion of the net benefits and harms for each guideline statement, the EP members rated each recommendation using the GRADE evidence-to-decision framework. This allows for a systematic way to document panel members’ judgement for each of the recommendations.[4]

Evidence-to-Decision Framework (EtD) Domains

Problem Priority	• Is the problem a priority and a recommendation is needed to address it? • Are there consequences that are serious if the problem is not addressed?

Benefits and Harms	• Are the desirable anticipated effects large? • Are the undesirable anticipated effects small? • Are the desirable effects large relative to undesirable effects?
Values and Preferences of Stakeholders	• Is there certainty of how stakeholders (patients, clinicians) value the outcomes? • Is there variability on how patients and clinicians value the outcomes? • Will there be different decisions from key stakeholders because of the different values placed on the outcomes?
Resources Required	• If the Recommendation is made, how large are the resource requirements?
Health Equity	• Are there groups or settings that might be disadvantaged in relation to the Recommendation being considered? • Are there different baseline conditions across groups or settings that affect the absolute effectiveness of the Recommendation or the importance of the problem for disadvantaged groups or settings? • Are there important considerations that should be made when implementing the Recommendation in order to ensure that inequities are reduced, if possible, and that they are not increased?
Feasibility	• Is the option (or recommendation) feasible to implement? • Is the Recommendation sustainable? Are there important barriers that are likely to limit the feasibility of implementing the Recommendation? If yes, do these barriers require consideration when implementing the Recommendation?
Acceptability	• Is the option acceptable to key stakeholders? • Are there key stakeholders that would not accept the distribution of the benefits, harms or costs? • Are there key stakeholders that would not accept the costs or undesirable effects in the short term for desirable effects (benefits) in the future?

Supplemental Table 3 provides a summary of the EP judgments within the EtD framework for each recommendation statement.

Assessing Quality and Risk of Bias

An assessment of the quality of the evidence was performed for all retained studies following application of the inclusion and exclusion criteria. Using this method, studies deemed be of low quality would not be excluded from the systematic review, but would be retained, and their methodological strengths and weaknesses discussed where relevant. To define an overall risk of bias rating for each included study, validated study-type specific tools were used to assess the risk of bias, plus additional important quality features were extracted. Specific details for each study type are outlined below.

- Systematic Reviews (SRs) and Meta-analyses were assessed as per the Assessing the Methodological Quality of Systematic Reviews (AMSTAR) [5] tool.

- Randomized Control Trials (RCTs) were assessed using the Cochrane Risk of Bias tool[6].

- Single-arm non-randomized phase I and II clinical trials (NRCTs), prospective cohort studies (PCS), prospective-retrospective cohort studies (PRCS), retrospective cohort studies (RCS), and case-control studies (CCS) were assessed using the Risk of Bias in Non-randomized Studies of Intervention (ROBINs-I)[7] tool.

In the following sections, the quantity of the evidence as determined by the number of studies that met our inclusion criteria and were retained, the evidence type as determined by study design, the quality of that evidence as determined by the quality assessment, and its consistency are all reported, both as individual studies and in totality, statement by statement. Definitions of the certainty of evidence is presented in Supplemental Table 4.

A total of 98 studies identified by the SR informed the recommendations. Although data was extracted from 121 studies, 23 studies contained insufficient detail to inform statements and 8 studies reported on outcomes that were not relevant to this guideline (Figure 2). The body of evidence was comprised of 2 SRs, 18 RCTs, 6 PCSs, and 72 RCSs. Risk of bias assessments for the systematic reviews, RCTs, prospective and retrospective cohort studies can be found in Supplemental Tables 5-8 respectively. The GRADE certainty of evidence for each outcome informing a recommendation and the overall certainty ranking for the statement is presented in Supplemental Table 9.

Open Comment Period and Organizational Review

An open comment period was held from March 31 to April 23, 2021, on the CAP web site (www.cap.org). Six draft statements, demographic questions, and questions to assess feasibility were posted for peer review. An announcement was sent to the following societies deemed to have interest.

Medical societies

- American Society for Clinical Oncology (ASCO)
- Association for Molecular Pathology (AMP)
- American Society for Clinical Pathology (ASCP)
- American College of Chest Physicians (CHEST)
- American College of Medical Genetics and Genomics (ACMG)
- American Society for Investigative Pathology (ASIP)
- American Society of Cytopathology (ASC)
- American Thoracic Society (ATS)
- Arthur Purdy Stout Society (APSS)
- Association of Community Cancer Centers (ACCC)
- Association of Directors of Anatomic and Surgical Pathology (ADASP)
- Association of Pathology Chairs (APC)
- British Thoracic Oncology Group
- Canadian Association of Pathologists (CAP-APC)
- European Society for Medical Oncology (ESMO)
- European Society of Thoracic Surgeons
- Indian Society for the Study of Lung Cancer
- International Association for the Study of Lung Cancer (IASLC)
- International Thoracic Oncology Nurses Forum
- Korean Association for the Study of Lung Cancer
- National Comprehensive Cancer Network (NCCN)
- National Lung Cancer Forum for Nurses
- Papanicolaou Society of Cytopathology (PSC)
- Pulmonary Pathology Society (PPS)
- Quality Initiative in Interpretive Pathology (QIIP) Canadian Partnership Against Cancer
- Russian Society of Clinical Oncology
- Sociedade Brasileira de Cirurgia Torácica (Brazilian Society of Thoracic Surgery)
- Sociedade Brasileira de Patologia (Brazilian Society of Pathology)
- Society to Improve Diagnoses in Medicine (SIDM)
- The Japan Lung Cancer Society

- United States & Canadian Academy of Pathology (USCAP)

Patient Advocacy Groups
- American Cancer Society
- American Lung Association
- Bonnie J. Addario Lung Cancer Foundation (ALCF)
- Cancer Leadership Council
- Cancer Research and Prevention Foundation
- Caring Ambassadors Lung Cancer Program
- Dusty Joy Foundation
- EX: Re-learn Live without Cigarettes
- Free Me From Lung Cancer
- Free to Breathe
- Global Lung Cancer Coalition
- Global Resource for Advancing Cancer Education
- International Thoracic Oncology Nursing Forum
- Lung Cancer Alliance
- Lung Cancer Foundation of American (LCFA)
- Lung Cancer Research Foundation (LCRF)
- LUNGevity Foundation
- Mesothelioma Applied Research Foundation
- My Cancer Genome
- Partnership Against Cancer American Cancer Society
- Prevent Cancer Foundation
- Roy Castle Lung Cancer Foundation
- UICC Global Cancer Control Community
- Union for International Cancer Control
- Uniting Against Lung Cancer
- Women Against Lung Cancer in Europe

Government and other stakeholders
- Canada Food and Drug Administration
- Centers for Medicare & Medicaid Services (CMS)
- Centers for Disease Control and Prevention (CDC)
- US Food and Drug Administration (FDA)
- Veteran's Affairs (VA) and Department of Defense (DOD)

"Agree" and "Disagree" responses were captured for every proposed recommendation. The EP reviewed all the comments. Resolution of all changes was obtained by majority consensus of the panel using a modified Delphi technique (discussion at an in-person meeting, rounds of teleconference webinars, email discussion and multiple edited recommendations) amongst the panel members. The final recommendations were approved by the EP with a formal vote. Neither formal cost analysis nor cost effectiveness models were performed.

Organizational review was instituted to review and approve the guideline. An independent review panel (IRP) representing the Council on Scientific Affairs was assembled to review and approve the guideline for the CAP. The IRP was masked to the expert panel and vetted through the COI process. Collaborating organizations were provided the guideline for approval. Once approved, the collaborating organizations' names were added to the guideline title as official collaborators.

Dissemination Plans

The CAP hosts a resource page which includes a link to the manuscript and supplement; a summary of the recommendations, a teaching PowerPoint (Microsoft Corporation, Redmond, WA), and a frequently asked question (FAQ) along with other additional tools such as webinar recordings as applicable. The guideline is promoted and presented at various society meetings and distributed to the societies listed in the peer review.

Recommendation Statements

Statement 1. In patients with advanced non-small cell lung cancer (NSCLC), pathologists should use a validated PD-L1 immunohistochemistryl (IHC) expression assay, in conjunction with other targetable genomic biomarker assays where appropriate, to optimize selection for treatment with immune checkpoint inhibitors (ICI).

The strength of recommendation is *strong*. The certainty of evidence to support this recommendation is *moderate*.

The evidence base for this statement includes 2 SRs [8, 9], 14 RCTs[10-23], 4 RCT post-hoc analyses[24-27], 4 PCSs [28-31], and 10 RCSs [32-41]. Identified studies reported on OS [8, 9, 11-18, 24-35, 37-44] and RR [10, 11, 14-17, 25-28, 30-34, 36, 37, 39-42] of various immunotherapy and ICI agents in tumors that were PD-L1 positive. The certainty of evidence was moderate for both outcomes of interest based on an aggregate serious risk of bias across studies informing both outcomes, but evidence was not further downgraded for any domain (Supplemental Table 9). In both outcomes there was inconsistency of results across studies; however, this variability was determined to be a consequence of differences in immunotherapy agents, PD-L1 expression cut-offs, and patient population, so evidence was not downgraded.

Based on the available evidence, EP members concluded that survival and response rates of ICI therapy were correlated with PD-L1 expression status. After discussions, the EP defined the benefits of PD-L1 expression detection using a validated IHC assay as moderate and the harms of this testing as small, and that the benefits thus outweighed the harms. It is expected that this guidance will be acceptable to key stakeholders and feasible to implement. Implementation of this guidance is expected to have no impact on health equity and the resource requirements were considered to be negligible (Supplemental Table 3).

Statement 2. Pathologists should ensure appropriate validation has been performed on all specimen types and fixatives.
Note: Specific validation requirements are out of scope with this guideline and laboratories should refer to the Principles of Analytic Validation of Immunohistochemical Assays Guideline[45] for details on how to validate IHC specimens.

The strength of recommendation is *conditional*. The certainty of evidence to support this guideline statement is *low*.

The statement is informed by 1 RCT post-hoc analysis[25], 2 PCSs [46, 47], and 32 retrospective studies[48-79]. Studies informing the statement reported on ICI therapy RRs and survival rates using various specimen types[25, 70], PD-L1 status concordance of various specimen types [47, 49-58, 60, 61, 63-69, 71, 73, 74, 76-79], diagnostic test characteristics of PD-L1 expression detection using various specimen types [49, 50, 60, 63, 76], PD-L1 status using multiple sample preparation types[46, 72], and both interobserver and intraobserver agreement for PD-L1 status using surgical sections[48, 57, 59, 62, 73], and cytology samples[57, 60, 73-75]. The certainty of evidence across the 19 outcomes ranged from very low through moderate (Supplemental Table 9). Assessment was based on an aggregate risk of bias of serious and very serious depending on the outcome, plus downgrading of evidence in some outcomes for inconsistency.

Based on the paucity of homogeneous evidence for any one sample type, the EP members concluded that PD-L1 expression was optimally determined using the best sample collected. After discussions, the EP defined the benefits of PD-L1 expression detection using the best sample available as moderate; however, the harms were also defined as moderate and the overall certainty of evidence was low, leading to the EP members to conclude that balance of effects probably favored testing in the best available sample. The EP also discussed that there was possibly important variability in the values and preferences of key stakeholders, but the guidance is expected to be acceptable and feasible to implement. Implementation of this guidance is expected to have no impact on health equity and the resource requirements were considered to be negligible (Supplemental Table 3).

Statement 3. When feasible, pathologists should use clinically validated PD-L1 IHC assays as intended.

The strength of recommendation is *conditional*. The certainty of evidence to support this guideline statement is *very low*.

The evidence base supporting this statement includes 2 systematic reviews[8, 9] , 12 RCTs[10-18, 42-44] , 4 RCT post-hoc analyses[24-27] , 4 PCSs[28-31] , and 29 RCSs [32-36, 38, 40, 41, 48, 69, 77, 80-97] . The included studies reported on PD-L1 status concordance[40, 41, 44, 48, 69, 77, 82-89, 92-94, 96, 98], diagnostic test characteristics[96] , and interobserver agreement[80, 81, 88, 90, 91, 94, 95, 97] of various combinations of clinically validated PD-L1 assays. Additionally, clinically trials and observational studies reporting on ICI therapy survival and response rates[8-18, 24-44] , leading to the clinically validation of the assays, were used as indirect evidence to support this statement. Certainty of evidence for the outcomes were assessed as low and very low based on an aggregate risk of bias ranging from serious to extremely serious across the outcomes, as well as further downgrading for inconsistency, imprecision, and indirectness in specific outcomes (Supplement Table 9).

The EP members discussed clinically validated versus laboratory developed PD-L1 IHC assays at length. Based on the available evidence, the EP members determined that use of a clinically validated assay carried moderate benefits; however, the harms of these assays, including availability of the staining platforms and clones, and specific training for each assay, were also defined as moderate. The EP concluded that the balance of effects did not favor clinically validated or laboratory developed assay. Further to this, EP members determined that use of clinically validated assays carried a large cost and could lead to reduced health equity. A conditional recommendation was based on the clinically validation of these assays and their established ability to predict immunotherapy response but with an understanding of the limitations of this guidance. The EP concluded that the guidance was probably be acceptable to key stakeholders and probably feasible to implement (Supplemental Table 3).

Statement 4. Pathologists that choose to use laboratory developed tests (LDTs) for PD-L1 expression should validate according to the requirement of their accrediting body.

The strength of recommendation is *strong*. The certainty of evidence to support this guideline statement is *very low*.

The evidence base informing this statement is comprised of nine RCSs [67, 91, 95, 99-104]. A study reported on ICI response rates using 22C3 and 73-10[104] , while other studies reported on PD-L1 status concordance[67, 91, 99-104] , diagnostic test characteristics[91] , and interobserver agreement[91, 95] of LDTs when compared with clinically validated IHC assays. Certainty of evidence for the outcomes of interest were assessed as low and very low (Supplemental Table 9). All outcomes assessed as low were informed by single studies with very serious risk of bias. Outcomes assessed as very low were supported by multiple studies, but studies were limited by very serious and extremely serious aggregate risk of bias, as well as inconsistency.

Although using a clinically validated assay is preferred, this is not always feasible, and many laboratories need to develop LDTs. Following discussions on the available evidence, the EP concluded that the benefits of validating all LDTs were large, while the harms of the validation were defined as trivial, leading to a balance of effects that favored validation. Although the EP determined that this guidance would require moderate resources, it likely would have no impact on health equity and would be acceptable to key stakeholders and feasible to implement (Supplemental Table 3). Given the low overall certainty of evidence when comparing PD-L1 status between clinically validated assays and the LTDs, the EP members concluded that not validating LDTs could lead to substantial harms to patients. Based on this assessment and other domains within the EtD, the EP drafted this recommendation as strong, despite a certainty of evidence rating of low.

Statement 5. Pathologists should report PD-L1 IHC results using a percent expression score.

The strength of recommendation is conditional. The certainty of evidence to support this guideline statement is very low.

The guideline statement is supported by 3 RCTs[15, 25, 42] and 7 RCSs [32, 34, 81, 88, 90, 91, 94]. Identified studies reported on ICI RRs[15, 25, 32, 34, 42] and survival rates[15, 25, 32, 34, 42] stratified by specific PD-L1 tumor proportion score (TPS) thresholds. Additional studies evaluated interobserver agreement using multiple IHC clones also stratified by TPS score[81, 88, 90, 91, 94]. The certainty of evidence for RRs was assessed as low based on a very serious aggregate risk of bias across the studies reporting on the outcome, but no further downgrading. For the other two outcomes, certainty was assessed as very low based on very serious risk of bias plus downgrading for inconsistency (Supplemental Table 9).

Based on the available evidence, the EP concluded that reporting PD-L1 expression as a percent score carried moderate benefits and only small harms, leading to the determination that benefits probably outweighed the potential harms. It is expected that this guidance will be acceptable to key stakeholders and feasible to implement. Implementation of this guidance is expected to have no impact on health equity and the resource requirements were considered to be negligible (Supplemental Table 3).

Statement 6. Clinicians should not use tumor mutation burden (TMB) alone to select patients with advanced NSCLC for immune checkpoint inhibitors based on insufficient evidence in this population.

The strength of recommendation is *conditional*. The certainty of evidence to support this guideline statement is *very low*.

The evidence base is comprised of eight RCSs reporting on ICI RRs[105, 106] and survival rates[32, 35, 38, 105-109] when correlated with TMB status. Certainty of evidence for RRs was assessed as very low based on extremely serious risk of bias in the studies reporting on this outcome but no further downgrading in any domain (Supplemental Table 9). For survival rates, the certainty of evidence was assessed as very low as a consequence of very serious aggregate risk of bias plus additional downgrading for inconsistency in the reported outcomes (Supplemental Table 9).

This conditional recommendation was based on the trivial benefits of using TMB to select patients for ICI therapies paired with the moderate harms of its use. The balance of effects favored not using TMB, as did the moderate costs and probable reduced health equity that would be associated with recommending its use. Further to these domains, the EP concluded that guidance in support of TMB would not be acceptable to key stakeholders and probably not be feasible to implement (Supplemental Table 3).

Supplemental Table 1. Key Questions and PICO Elements

Pre-Analytical Stage		
KQ1a. In early stage, resectable NSCLC patients who are being considered for immune checkpoint inhibitors (ICI), does PD-L1 and tumor mutation burden (TMB) testing improve treatment response rates and survival rates?		
Population		
Patients with early stage unresectable NSCLC		
Intervention	**Comparator**	**Outcomes**
PD-L1 testing TMB testing	• Different PD-L1 test • TMB test • Not testing • Single arm	Critical • Response rates • Survival rates
KQ1b. In early stage NSCLC patients who are being considered for ICI , does PD-L1 and TMB testing improve treatment response rates and survival rates?		
Population		
Patients with early stage unresectable NSCLC		
Intervention	**Comparator**	**Outcomes**
PD-L1 testing TMB testing	• Different PD-L1 test • TMB test • Not testing • Single arm	Critical • Response rates • Survival rates
KQ1c. In advanced stage NSCLC patients being considered for immune checkpoint inhibitors , does PD-L1 and TMB testing improve treatment response rates and survival?		
Population		
Patients with advanced stage NSCLC		
Intervention	**Comparator**	**Outcomes**
PD-L1 testing TMB testing	• Different PD-L1 test • TMB test • Not testing • Single arm	Critical • Response rates • Survival rates
KQ2. When selecting patients for anti-PD1 and anti-PD-L1 therapy, does testing of different specimen types provide concordant clinical outcomes?		
Population		
Specimens from patients with early and advanced stage NSCLC being considered for immune checkpoint inhibitors		
Intervention	**Comparator**	**Outcomes**
Primary tumor samples Metastatic samples Cytology samples (smears, liquid-based, and blocks) Core needle biopsy samples Resection/surgical specimens Archived tissue Tissue microarray	• Any included intervention • Single arm	Critical • Response rates • Survival rates • Clinical validity • Sample concordance A. Important • Tissue utilization
KQ3. Does the use of immunotherapy in advanced NSCLC patients with targetable *ALK*, *EGFR*, *ROS1*, or *BRAF* molecular alterations affect their long-term clinical outcomes?		
Population		
Patients with early and advanced stage NSCLC being considered for immune checkpoint inhibitors		
Intervention	**Comparator**	**Outcomes**

Anti-PD1/PD-L1 therapy	Standard of care (includes targeted therapy and chemotherapy)	Critical • Response rates • Survival rates • Adverse events/toxicity
Pre-Analytical and Analytical Stage		
KQ4. When selecting patients for anti-PD1 and anti-PD-L1 therapy, does TMB testing have the analytical validity to identify a complementary population who will benefit from therapy?		
Population		
Patients with early and advanced stage NSCLC being considered for immune checkpoint inhibitors		
Intervention	**Comparator**	**Outcomes**
TMB testing Combination TMB and PD-L1 testing	• PD-L1 testing • Single arm • No testing *Note:* Gold standard defined by the study	Critical • Response rates • Survival rates • Diagnostic test characteristics • Analytical validity • Assay concordance
KQ5a. In NSCLC patients with more than one available sample, do multiple samples from the same site provide concordant PD-L1 and TMB testing results and downstream clinical outcomes?		
Population:		
Multiple samples from patients with early and advanced stage NSCLC cancer being considered for immune checkpoint inhibitors		
Intervention	**Comparator**	**Outcomes**
Testing of secondary/additional sample(s) B.	• Testing of primary samples • Single arm	Critical • Response rates • Survival rates • Diagnostic test characteristics • Clinical and analytical validity • Sample concordance Important • Tissue utilization
KQ5b. In NSCLC patients with more than one available sample, do multiple samples from different sites provide concordant PD-L1 and TMB testing results and downstream clinical outcomes?		
Population:		
Multiple samples from patients with early and advanced NSCLC cancer being considered for immune checkpoint inhibitors		
Intervention	**Comparator**	**Outcomes**
Testing of secondary, archived or metastatic sample	• Testing of primary samples • Single arm	Critical • Response rates • Survival rates • Diagnostic test characteristics • Clinical and analytical validity • Sample concordance Important • Tissue utilization
KQ6a. Does clinical validity of PD-L1 testing differ by levels of PD-L1 expression in tumor cells?		
Population:		
Specimens from early and advanced stage NSCLC patients being considered for immune checkpoint inhibitors		
Intervention	**Comparator**	**Outcomes**
Negative expression level Positive expression level	• Any included intervention • Single arm	Critical • Response rates • Survival rates

Study defined expression level cut-off Tumor cell score		• Diagnostic test characteristics • Clinical and analytical validity
KQ6b. Does addition of PD-L1 expression levels in immune cells to expression in tumor cells improve patient clinical outcomes?		
Population:		
Specimens from early and advanced stage NSCLC patients being considered for immune checkpoint inhibitors		
Negative expression level Positive expression level Study defined expression level cut-off Immune cell score	• Any included intervention • Single arm	Critical • Response rates • Survival rates • Diagnostic test characteristics • Clinical and analytical validity
Analytical Stage		
KQ7. How reproducible are PD-L1 tumor cell scores and immune cell scores across specimen types?		
Population:		
Specimens from early and advanced stage NSCLC patients		
Intervention	**Comparator**	**Outcomes**
Primary tumor samples Metastatic samples Cytology samples (smears, liquid-based, and blocks) Core needle biopsy samples Resection/surgical specimens Archived/saved tissue samples Tissue microarray	• Any included intervention • Single arm	Critical • Diagnostic test characteristics • Analytical validity ○ Includes interobserver agreement • Sample concordance Important • Tissue utilization C.
KQ8. Do the available PD-L1 assays provide concordant expression profiles when evaluating the same sample and which IHC expression cut-off provides the most reproducible expression categorization across the assays?		
Population:		
Specimens from early and advanced stage NSCLC patients		
Intervention: PD-L1 Antibody Clones		
• QR1 • E1L3N	• 73-10 • 22C3	• SP142 • 28-8	• SP263 D.
Intervention: Expression Level Cut-Offs	**Comparator**	**Outcomes**
H-score Study defined expression level cut-off Ratio of tumor cell PD-L1 to immune cell PD-L1	• Any other expression level cut-off • 22c3 companion dx – ideal/gold standard but flawed • No testing/single arm • TMB testing *Note:* Gold standard defined by the study	Critical • Diagnostic test characteristics • Analytical validity ○ Includes interobserver agreement • Assay concordance

Patient Population	All adult early stage and advanced stage NSCLC patients being considered for immune checkpoint inhibitors • For KQs focusing on analytical outcomes, studies that enroll NSCLC patients undergoing testing without detail on subsequent immuno-oncology therapy will be considered for inclusion • For KQs focusing on clinical outcomes, PD-L1 and TMB studies without testing/assay specific details will be considered for inclusion
Setting	Academic and community laboratory settings
Minimum Sample Size	30 patients per study arm
Search Dates	2010 – 2021
Included Study Types	• Guidelines • Systematic reviews with and without meta-analysis • Randomized controlled trials • Comparative and single-arm observational studies with prospective or retrospective design • Case-control studies

Abbreviations: ALK, Anaplastic Lymphoma Kinase; BRAF, B-Raf Proto-Oncogene; EGFR, Epidermal Growth Factor Receptor; KQ, key question; NSCLC, Non small cell lung cancer; PICO, population, intervention, comparator, outcomes; PD-L1, programmed dealth ligand-1; ROS1, ROS Proto-Oncogene 1; TMB, tumor mutation burden

Supplemental Table 2: Grades for Strength of Recommendations

Designation	Recommendation	Evidence to Decision (EtD) Judgement
Strong Recommendation	Recommend for or against a particular practice (Can include "must" or "should")	Supported by assessment with the GRADE EtD framework showing expert panel (EP) consensus of judgements directed to the far right or far left poles of the framework
Conditional Recommendation	Recommend for or against a particular practice (Can include "should" or "may")	Supported by assessment with the GRADE EtD framework showing EP consensus of judgements directed towards the center of the framework or with a dispersed pattern

Derived from Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group materials[4, 110]

Supplemental Table 3. Evidence-to-Decision Framework

Statement 1. In patients with advanced non-small cell lung cancer (NSCLC), pathologists should use a validated PD-L1 immunohistochemical (IHC) expression assay, in conjunction with other targetable genomic biomarker assays where appropriate, to optimize selection for treatment with immune checkpoint inhibitors (ICI).

Summary of Judgements

Criteria	*Favors the comparison*		*Neutral*	*Favors the intervention*	
Problem	No	Probably no		Probably yes	**Yes +**
Desirable Effects	Trivial	Small		**Moderate +**	Large
Undesirable Effects	Large	Moderate		**Small +**	Trivial
Certainty of Effects	Very low	Low		**Moderate +**	Large
Values	Important certainty or variability	Possibly important uncertainty or variability		Probably no important uncertainty of variability	**No important uncertainty of variability +**
Balance of Effects	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	**Favors the intervention +**
Resources Required	Large costs	Moderate costs	**Negligible costs and savings +**	Moderate savings	Large savings
Equity	Reduced	Probably reduced	**Probably no impact +**	Probably increased	Increased
Acceptability	No	Probably no		Probably yes	**Yes +**
Feasibility	No	Probably no		Probably yes	**Yes +**

Statement 2. Pathologists should ensure appropriate validation has been performed on all specimen types and fixatives.
Note: Specific validation requirements are out of scope with this guideline and laboratories should refer to the *Principles of Analytic Validation of Immunohistochemical Assays Guideline*[45] for details on how to validate IHC specimens.

Summary of Judgements

Criteria	*Favors the comparison*		*Neutral*	*Favors the intervention*	
Problem	No	Probably no		Probably yes	**Yes +**
Desirable Effects	Trivial	Small		**Moderate +**	Large
Undesirable Effects	Large	**Moderate +**		Small	Trivial
Certainty of Effects	Very low	**Low +**		Moderate	Large
Values	Important certainty or variability	**Possibly important uncertainty or variability +**		Probably no important uncertainty of variability	No important uncertainty of variability
Balance of Effects	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	**Probably favors the intervention +**	Favors the intervention

Resources Required	Large costs	Moderate costs	**Negligible costs and savings +**	Moderate savings	Large savings
Equity	Reduced	Probably reduced	**Probably no impact +**	Probably increased	Increased
Acceptability	No	Probably no		Probably yes	**Yes +**
Feasibility	No	Probably no		Probably yes	**Yes +**

Statement 3. When feasible, pathologists should use clinically validated PD-L1 IHC assays as intended.

Summary of Judgements

Criteria	*Favors the comparison*		*Neutral*	*Favors the intervention*	
Problem	No	Probably no		Probably yes	**Yes +**
Desirable Effects	Trivial	Small		**Moderate +**	Large
Undesirable Effects	Large	**Moderate +**		Small	Trivial
Certainty of Effects	Very low	**Low +**		Moderate	Large
Values	Important certainty or variability	**Possibly important uncertainty or variability +**		Probably no important uncertainty of variability	No important uncertainty of variability
Balance of Effects	Favors the comparison	Probably favors the comparison	**Does not favor either the intervention or the comparison +**	Probably favors the intervention	Favors the intervention
Resources Required	**Large costs +**	Moderate costs	Negligible costs and savings	Moderate savings	Large savings
Equity	Reduced	**Probably reduced +**	Probably no impact	Probably increased	Increased
Acceptability	No	Probably no		**Probably yes +**	Yes
Feasibility	No	Probably no		**Probably yes +**	Yes

Statement 4 . Pathologists that choose to use laboratory developed tests (LDTs) for PD-L1 expression should validate according to the requirement of their accrediting body.

Summary of Judgements

Criteria	*Favors the comparison*		*Neutral*	*Favors the intervention*	
Problem	No	Probably no		Probably yes	**Yes +**
Desirable Effects	Trivial	Small		Moderate	**Large +**
Undesirable Effects	Large	Moderate		Small +	**Trivial +**
Certainty of Effects	Very low	**Low +**		Moderate	Large
Values	Important certainty or variability	**Possibly important uncertainty or variability +**		Probably no important uncertainty of variability	No important uncertainty of variability
Balance of Effects	Favors the comparison	Probably favors the comparison	Does not favor either the	Probably favors the intervention	**Favors the intervention**

			intervention or the comparison		+
Resources Required	Large costs	**Moderate costs +**	Negligible costs and savings	Moderate savings	Large savings
Equity	Reduced	Probably reduced	**Probably no impact +**	Probably increased	Increased
Acceptability	No	Probably no		Probably yes	**Yes +**
Feasibility	No	Probably no		Probably yes	**Yes +**

Statement 5. Pathologists should report PD-L1 IHC results using a percent expression score.

Summary of Judgements

Criteria	*Favors the comparison*		*Neutral*	*Favors the intervention*	
Problem	No	Probably no		Probably yes	**Yes +**
Desirable Effects	Trivial	Small		**Moderate +**	Large
Undesirable Effects	Large	Moderate		**Small +**	Trivial
Certainty of Effects	**Very low +**	Low		Moderate	Large
Values	Important certainty or variability	**Possibly important uncertainty or variability +**		Probably no important uncertainty of variability	No important uncertainty of variability
Balance of Effects	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	**Probably favors the intervention +**	Favors the intervention
Resources Required	Large costs	Moderate costs	**Negligible costs and savings +**	Moderate savings	Large savings
Equity	Reduced	Probably reduced	**Probably no impact +**	Probably increased	Increased
Acceptability	No	Probably no		Probably yes	**Yes +**
Feasibility	No	Probably no		Probably yes	**Yes +**

Statement 6. Clinicians should not use tumor mutation burden (TMB) alone to select patients with advanced NSCLC for immune checkpoint blockade based on insufficient evidence in this population.

Summary of Judgements

Criteria	*Favors the comparison*		*Neutral*	*Favors the intervention*	
Problem	**No +**	Probably no		Probably yes	Yes
Desirable Effects	**Trivial +**	Small		Moderate	Large
Undesirable Effects	Large	**Moderate +**		Small	Trivial
Certainty of Effects	Very low	**Low +**		Moderate	Large
Values	Important certainty or variability	Possibly important uncertainty or variability		Probably no important uncertainty of variability	**No important uncertainty of variability +**
Balance of Effects	**Favors the comparison +**	Probably favors the comparison	Does not favor either the	Probably favors the intervention	Favors the intervention

			intervention or the comparison		
Resources Required	Large costs	**Moderate costs +**	Negligible costs and savings	Moderate savings	Large savings
Equity	Reduced	**Probably reduced +**	Probably no impact	Probably increased	Increased
Acceptability	**No +**	Probably no		Probably yes	Yes
Feasibility	No	**Probably no +**		Probably yes	Yes

Abbreviations: PD-L1, programmed death ligand-1

Supplemental Table 4: Certainty of Evidence	
Designation	**Description**
High	There is high confidence that available evidence reflects true effect. Further research is very unlikely to change the confidence in the estimate of effect. Included studies will be of high or intermediate quality.
Moderate	There is moderate confidence that available evidence reflects true effect. Further research is likely to have an important impact on the confidence in estimate of effect and may change the estimate. Included studies will be of intermediate or low quality.
Low	There is limited confidence in the estimate of effect. The true effect may be substantially different from the estimate of the effect. Included studies will be of low quality.
Very Low	There is very little confidence in the estimate of effect. The true effect is likely to be substantially different from the estimate of effect. Any estimate of effect is very uncertain. Included studies will be of low or very low quality.

Derived from Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group Materials. [110]

Supplemental Table 5. Risk of Bias Assessment of Included Systematic Reviews

Study		Cao et al[8] 2019	Kim et al[9] 2019
AMSTAR Assessment[5]	A priori design	No	Yes
	Duplicate study selection and data extraction	No	CA
	Comprehensive literature search	CA	Yes
	Publication status as inclusion criterion	No	No
	List of included and excluded studies	Yes	Yes
	Characteristics of included studies	Yes	Yes
	Study quality assessment conducted	Yes	Yes
	Quality assessment used in formulating conclusions	No	Yes
	Appropriate methods to combine findings	Yes	Yes
	Publication bias assessment	Yes	Yes
	Conflict of interest reported	Yes	Yes
Reported funding sources		Yes	Yes

Abbreviations: AMSTAR, A MeaSurement Tool to Assess systematic Reviews; CA, can't answer.

Supplemental Table 6. Risk of Bias Assessment of Included Randomized Controlled Trials

Study	Cochrane Risk of Bias Tool Assessment							Additional Quality Features			
	Random sequence generation	Allocation concealment	Blinding – patients and conductors	Blinding – outcome assessors	Complete outcome data	Selective outcome reporting	Overall Risk of Bias	Validated and reliable measures	Adequately powered	Reported funding sources	Industry funded
Park et al, [21] 2021	LR	HR	HR	HR	LR	LR	Int	Y	Y	Y	Y
Borghaie et al, [10] 2015	HR	HR	HR	HR	LR	LR	High	Y	Y	Y	Y
Brahmer et al, [11] 2015	UR	UR	UR	UR	LR	HR	Int	Y	NS	Y	Y
Fehrenbacher, [12] 2016	LR	HR	HR	HR	LR	LR	Int	Y	Y	Y	Y
Garon et al, [13] 2019	UR	HR	HR	HR	LR	LR	Int	Y	Y	Y	Y
Garon et al, [24] 2015	UR	HR	HR	HR	LR	LR	Int	Y	Y	Y	Y
Hellmann et al, [14] 2019	UR	UR	UR	UR	LR	LR	Int	Y	Y	Y	Y
Herbst et al, [23] 2021	LR	HR	HR	LR	LR	LR	Int	Y	Y	Y	Y
Herbst et al, [25] 2019	UR	HR	HR	HR	LR	HR	Int	Y	NS	Y	Y
Horn et al, [26] 2018	HR	HR	UR	UR	LR	LR	Int	Y	NS	Y	Y
Hui et al, [27]2017	UR	UR	HR	HR	LR	LR	Int	Y	Y	Y	Y
Mok et al, [15] 2019	LR	HR	HR	HR	LR	LR	Int	Y	Y	Y	Y
Ready et al, [16] 2019	HR	HR	HR	HR	LR	HR	High	Y	NS	Y	Y
Rittmeyer et al, [17] 2017	LR	HR	HR	HR	LR	LR	Int	Y	Y	Y	Y
West et al, [18] 2019	LR	UR	HR	HR	LR	LR	Int	Y	Y	Y	Y
Jassem et al, [22] 2021	UR	UR	UR	UR	LR	LR	Int	Y	Y	Y	Y
Reck et al, [19] 2021	UR	HR	HR	HR	HR	LR	High	Y	Y	Y	Y
Paz-Ares et al, [20] 2022	UR	UR	HR	HR	LR	LR	Int	Y	Y	Y	Y

Abbreviations: HR, high risk; Int, intermediate; LR, low risk; N, no; NS, no statistical analysis: UR, unclear risk; Y, yes.

Supplemental Table 7. Risk of Bias Assessment of Included Prospective Cohort Studies

Study	ROBINS-I Assessment								Additional Quality Features		
	Confounding	Patient selection	Intervention classification	Deviation from intended intervention	Missing data	Outcome measurements	Selection of reported outcomes	Overall Risk of Bias	Adequately powered	Reported funding sources	Industry funded
Antonia et al, [28] 2019	MR	MR	LR	LR	MR	MR	MR	MR	NS	Y	N
Gettinger et al, [29] 2018	LR	LR	LR	LR	MR	MR	LR	MR	NS	N	U
Gettinger et al, [30] 2016	MR	MR	LR	LR	MR	SR	MR	SR	NS	Y	Y
Peters et al, [31] 2017	MR	MR	LR	LR	LR	MR	LR	MR	NS	Y	Y
Vigiliar et al, [46] 2019	LR	MR	LR	LR	LR	MR	LR	MR	Y	Y	N
Wang et al, [47] 2019	MR	MR	LR	LR	MR	MR	LR	MR	Y	Y	N

Abbreviations: CR, critical risk; LR, low risk; MR, moderate risk; N, no; NS, no statistical analysis: ROBINS-I, Risk Of Bias In Non-Randomized Studies - of Interventions;SR, serious risk; U, unclear; Y, yes.

Supplemental Table 8. Risk of Bias Assessment of Included Retrospective Cohort Studies

Study	ROBINS-I Assessment								Additional Quality Features		
	Confounding	Patient selection	Intervention classification	Deviation from intended intervention	Missing data	Outcome measurements	Selection of reported outcomes	Overall Risk of Bias	Adequately powered	Reported funding sources	Industry funded
Aguilar et al, [32] 2019	MR	CR	LR	LR	MR	MR	LR	CR	Y	Y	N
Ahn et al, [33] 2019	MR	CR	LR	LR	MR	LR	MR	CR	U	Y	N
Brunnstrom et al, [80] 2017	MR	SR	LR	LR	MR	MR	MR	SR	Y	Y	N
Chan et al, [48] 2018	MR	CR	LR	LR	MR	MR	LR	CR	NS	Y	N
Cooper et al, [81] 2017	MR	CR	LR	LR	LR	MR	LR	CR	NS	Y	Y
Edahiro et al, [34] 2019	MR	CR	LR	LR	MR	LR	LR	CR	Y	Y	N
Elfving et al, [49] 2019	MR	CR	LR	LR	MR	LR	LR	CR	NS	Y	N
Gradecki et al, [50] 2018	MR	CR	LR	LR	LR	MR	MR	CR	NS	Y	N
Grosu et al, [51] 2019	MR	CR	LR	LR	MR	LR	LR	CR	NS	Y	N
Hernandez et al, [52] 2019	MR	CR	LR	LR	LR	LR	LR	CR	Y	N	N
Ilie et al, [53] 2016	MR	CR	LR	LR	MR	MR	LR	CR	NS	Y	N
Keller et al, [54] 2018	MR	CR	LR	LR	MR	LR	MR	CR	Y	N	U
Kim et al, [55] 2017	MR	CR	LR	SR	MR	LR	LR	CR	Y	Y	N
Kim et al, [35] 2019	MR	CR	LR	MR	SR	MR	LR	CR	Y	Y	N
Kim et al, [56] 2017	MR	CR	LR	LR	LR	SR	MR	CR	NS	Y	N
Krawczyk et al, [82] 2017	MR	CR	LR	LR	LR	LR	LR	CR	N	Y	N
Kuempers et al[57] 2019	MR	CR	LR	LR	MR	MR	LR	CR	NS	N	U
Lin et al, [36] 2018	MR	CR	LR	LR	MR	LR	LR	CR	N	N	U
Liu et al, [107] 2019	MR	CR	LR	LR	MR	MR	LR	CR	Y	Y	N
Mei et al, [58] 2019	MR	CR	LR	LR	LR	LR	LR	CR	NS	Y	N
Munari et al, [99] 2018	MR	CR	LR	LR	MR	MR	LR	CR	NS	Y	N
Munari et al, [59] 2018	MR	CR	LR	LR	MR	MR	MR	CR	NS	Y	N
Munari et al, [60] 2019	MR	CR	LR	LR	MR	LR	LR	CR	NS	Y	N
Noll et al, [61] 2018	MR	CR	LR	LR	LR	MR	LR	CR	NS	Y	N
Oya et al, [37] 2017	MR	CR	LR	LR	MR	LR	MR	CR	Y	Y	N
Pang et al, [83] 2018	MR	CR	LR	LR	MR	MR	LR	CR	NS	Y	N
Rehman et al, [62] 2017	MR	CR	LR	LR	MR	MR	MR	CR	NS	Y	N
Rizvi et al, [38] 2018	MR	CR	LR	MR	MR	MR	LR	CR	Y	N	U
Saito et al, [84] 2018	MR	CR	LR	LR	LR	MR	MR	CR	NS	Y	Y
Sakata et al, [63] 2018	MR	CR	LR	LR	MR	LR	LR	CR	NS	Y	N
Sheffield et al, [64] 2016	MR	CR	LR	LR	LR	MR	LR	CR	Y	Y	N
Song et al, [85] 2019	MR	CR	LR	LR	MR	MR	LR	CR	NS	Y	N

Tamiya et al, [39] 2019	MR	CR	LR	LR	MR	MR	LR	CR	Y	Y	N
Teglasi et al, [65] 2019	MR	CR	LR	LR	MR	MR	LR	CR	Y	Y	N
Uruga et al, [66] 2017	MR	CR	LR	LR	MR	MR	MR	CR	N	Y	N
Villaruz et al, [100] 2019	MR	CR	LR	LR	MR	LR	LR	CR	Y	Y	N
Yeo et al, [86] 2017	MR	CR	LR	LR	MR	MR	LR	CR	N	Y	N
Adam et al, [101] 2018	MR	CR	LR	LR	LR	MR	LR	CR	NS	Y	Y
Beck et al, [87] 2019	MR	CR	LR	LR	LR	MR	LR	CR	NS	N	U
Chang et al, [108] 2019	MR	CR	LR	MR	MR	MR	MR	CR	NS	Y	Y
Fujimoto et al, [40] 2018	MR	CR	LR	MR	MR	MR	MR	CR	NS	N	U
Fujimoto et al, [88] 2018	MR	CR	LR	LR	LR	LR	LR	CR	NS	N	U
Hirsch et al, [98] 2017	MR	CR	LR	LR	SR	MR	MR	CR	NS	Y	Y
Humphries et al, [89] 2019	MR	SR	LR	MR	MR	MR	LR	SR	NS	Y	N
Illie et al, [67] 2018	MR	CR	LR	LR	MR	MR	MR	CR	NS	Y	Y
Kim et al, [102] 2017	MR	CR	LR	MR	SR	MR	MR	CR	NS	Y	N
Marchetti et al, [90] 2017	MR	CR	CR	LR	LR	MR	MR	CR	NS	Y	N
Munari et al, [91] 2019	MR	CR	LR	LR	MR	LR	LR	CR	Y	Y	N
Munari et al, [68] 2018	MR	CR	CR	LR	LR	SR	LR	CR	N	Y	N
Park et al, [92] 2019	MR	CR	LR	LR	LR	LR	MR	CR	NS	Y	N
Parra et al, [93] 2018	MR	CR	LR	MR	SR	MR	MR	CR	NS	Y	N
Ratcliffe et al, [94] 2017	MR	CR	LR	LR	LR	MR	MR	CR	NS	Y	Y
Rimm et al, [95] 2017	MR	CR	LR	LR	SR	MR	LR	CR	Y	Y	Y
Singal et al, [105] 2019	MR	CR	LR	MR	SR	MR	MR	CR	NS	Y	Y
Skov et al, [69] 2017	MR	CR	LR	LR	MR	LR	LR	CR	NS	Y	Y
Sughayer et al, [103] 2019	MR	CR	LR	LR	SR	MR	LR	CR	NS	N	U
Torous et al, [70] 2018	MR	CR	LR	LR	MR	MR	MR	CR	Y	Y	N
Tseng et al, [41]2018	MR	CR	LR	LR	MR	LR	MR	CR	Y	Y	N
Wang et al, [71] 2019	MR	CR	LR	LR	MR	LR	LR	CR	N	N	U
Wang et al, [72] 2018	MR	CR	LR	LR	LR	MR	LR	CR	Y	Y	N
Xu et al, [96] 2017	MR	CR	LR	LR	MR	MR	LR	CR	Y	Y	N
Tsao et al, [97] 2018	MR	CR	LR	LR	MR	MR	MR	CR	Y	Y	Y
Russell-Goldman et al, [73] 2018	MR	CR	LR	LR	LR	MR	LR	CR	Y	Y	N
Aggarwal et al, [106] 2020	MR	CR	LR	LR	MR	LR	MR	CR	N	Y	Y
Daverio et al, [74] 2020	MR	CR	LR	LR	LR	LR	LR	CR	N	Y	N
Grote et al, [104] 2020	MR	CR	LR	LR	MR	LR	LR	CR	NS	Y	Y
Hernandez et al, [75] 2020	MR	CR	LR	LR	MR	LR	MR	CR	N	N	U
Lou et al, [76] 2020	MR	CR	LR	LR	LR	LR	MR	CR	NS	Y	N
Song et al, [78] 2020	MR	CR	LR	LR	LR	LR	LR	CR	N	Y	N
Zou et al, [79] 2020	MR	CR	LR	LR	LR	LR	LR	CR	Y	Y	N
Suzuki et al, [109] 2022	MR	CR	LR	MR	MR	MR	MR	CR	NS	Y	N

Supplemental Table 9. GRADE Certainty of Evidence Assessment

Number of Studies and Design	Aggregate Risk of Bias	Inconsistency	Indirectness	Imprecision	Other [A]	Certainty of Evidence Grade for Outcome	Overall Certainty of Evidence Grade for Statement
STATEMENT 1							
Survival, ALL immunotherapy agents (critical outcome [B])							Moderate
2 SR, 13 RCT, 4 RCT post-hoc, 4 PCS, 9 RCS	Serious	Not serious [C]	Not serious	Not serious	None	Moderate	
Response Rates, ALL immunotherapy agents (critical outcome [B])							
8 RCT, 3 RCT post-hoc, 3 PCS, 8 RCS	Serious	Not serious [C]	Not serious	Not serious	None	Moderate	
STATEMENT 2							
Response Rate – Cytology Blocks and Surgical Resections (critical outcome [B])							Low
1 RCT post-hoc	Serious	Not serious [D]	Not serious	Not serious	Other	Moderate	
Response Rate – Archived and Fresh Samples (critical outcome [B])							
1 RCT post-hoc	Serious	Not serious [D]	Not serious	Not serious	Other	Moderate	
Survival Rate – Archived and Fresh Samples (critical outcome [B])							
1 RCT post-hoc	Serious	Not serious [D]	Not serious	Not serious	Other	Moderate	
PD-L1 Status Concordance – Primary and Metastatic Samples (critical outcome [B])							
7 RCS	Very Serious	Serious	Not serious	Not serious	None	Very Low	
PD-L1 Status Concordance – Cytology and Histology Samples (critical outcome [B])							
12 RCS	Very Serious	Not serious	Not serious	Not serious	None	Low	
Diagnostic Test Characteristics – Cytology Smear, Whole Section Reference Standard (critical outcome [B])							
1 RCS	Very Serious	Not serious [D]	Not serious	Not serious	None	Low	
Diagnostic Test Characteristics – Cytology Cell Blocks, Surgical Section Reference Standard (critical outcome [B])							
1 RCS	Very Serious	Not serious [D]	Not serious	Not serious	None	Low	
PD-L1 Status Concordance – Fluid Cell Blocks and Biopsy Specimens (critical outcome [B])							
2 RCS	Very Serious	Not serious	Not serious	Not serious	None	Low	
PD-L1 Status Concordance – Bronchial Biopsies and Surgical Resections (critical outcome [B])							
1 RCS	Very Serious	Not serious [D]	Not serious	Not serious	None	Low	
Diagnostic Test Characteristics – CNB Samples, Surgical Resection Reference Standard (critical outcome [B])							
2 RCS	Very Serious	Not serious	Not serious	Not serious	None	Low	
PD-L1 Status Concordance – CNB Samples and Resection Specimens (critical outcome [B])							
2 RCS	Very Serious	Not serious	Not serious	Not serious	None	Low	
Diagnostic Test Characteristics – EBUS-TBNA, Surgical Resections Reference Standard (critical outcome [B])							
1 RCS	Very serious	Not serious [D]	Not serious	Not serious	None	Low	

PD-L1 Status Concordance – EBUS-TBNA/FNA and Surgical Resections (critical outcome [B])							
1 PCS, 1 RCS	Very serious	Not serious	Not serious	Not serious	None	Low	
Interobserver Agreement – Surgical Resections (critical outcome [B])							
4 RCS	Very Serious	Not serious [E]	Not serious	Not serious	None	Low	
Interobserver Agreement – Cytology Specimens (critical outcome [B])							
5 RCS	Very Serious	Serious	Not serious	Not serious	None	Very Low	
Intraobserver Agreement – Cytology Specimens (critical outcome [B])							
1 RCS	Very Serious	Not serious [D]	Not serious	Not serious	None	Low	
Interobserver Agreement – Archived Samples (critical outcome [B])							
1 RCS	Very Serious	Not serious [D]	Not serious	Not serious	None	Low	
Intraobserver Agreement – Archived Samples (critical outcome [B])							
1 RCS	Very Serious	Not serious [D]	Not serious	Not serious	None	Low	
Raw PD-L1 Status – Multiple Sample Preparation Types (important outcome [B])							
1 PCS, 1 RCS	Very Serious	Serious	Not serious	Not serious	None	Very Low	
STATEMENT 3							
PD-L1 Status Concordance – 22C3 and SP142 (critical outcome [B])							Very Low
2 RCS	Very Serious	Not serious [F]	Not serious	Not serious	None	Low	
Diagnostic Test Characteristics – SP142, 22C3 Reference Standard (critical outcome [B])							
1 RCS	Very Serious	Not serious [D]	Not serious	Not serious	None	Low	
PD-L1 Status Concordance – 22C3 and 28-8 (critical outcome [B])							
2 RCS	Very Serious	Not serious	Not serious	Serious	None	Very Low	
PD-L1 Status Concordance – 22C3 and SP263 (critical outcome [B])							
7 RCS	Very Serious	Not serious	Not serious	Not serious	None	Low	
PD-L1 Status Concordance – 22C3, SP142, SP263, and 28-8 (critical outcome [B])							
4 RCS	Very Serious	Not serious	Not serious	Serious	None	Very Low	
PD-L1 Status Concordance – SP142 and SP263 (critical outcome [B])							
1 RCS	Very Serious	Not serious [D]	Not serious	Not serious	None	Low	
PD-L1 Status Concordance – 22C3, SP142, and SP263 (critical outcome [B])							
2 RCS	Very Serious	Not serious [F]	Not serious	Not serious	None	Low	
Interobserver Agreement – Multiple IHC Clones Stratified by TPS (critical outcome [B])							
5 RCS	Very Serious	Serious	Not serious	Not serious	None	Very Low	
Interobserver Agreement – Multiple IHC Clones (critical outcome [B])							
3 RCS	Very Serious	Not serious	Not serious	Not serious	None	Low	
Raw PD-L1 Status – Multiple IHC Clones (important outcome [B])							
1 RCS	Extremely Serious	Not serious [D]	Not serious	Not serious	None	Very Low	
Survival, ALL immunotherapy agents Survival, ALL immunotherapy agents (critical outcome [B])							
2 MA, 9 RCT, 3 RCT post-hoc, 6 PCS, 9 RCS	Serious	Not serious [F]	Serious [G]	Not serious	None	Low	
Response Rates, ALL immunotherapy agents (critical outcome [B])							

6 RCT, 3 RCT post-hoc, 4 PCS, 8 RCS	Serious	Not serious [F]	Serious [G]	Not serious	None	Low	
STATEMENT 4							
Immunotherapy Response Rates in Cytology Cell Blocks and Surgical Resections based on 22C3 and 73-10 Clones (critical outcome [B])							Very Low
1 RCS	Very Serious	Not serious [D]	Not serious	Not serious	None	Low	
PD-L1 Status Concordance – 22C3 LDT and 22C3 or SP263 (critical outcome [B])							
4 RCS	Very serious	Serious	Not serious	Not serious	None	Very Low	
PD-L1 Status Concordance – E1L3N LDT and 22C3 or SP142 or SP263 (critical outcome [B])							
2 RCS	Extremely Serious	Not serious	Not serious	Not serious	None	Very Low	
Diagnostic Test Characteristics – E1L3N LDT, SP263 or 22C3 Reference Standard (critical outcome [B])							
1 RCS	Very Serious	Not serious [D]	Not serious	Not serious	None	Low	
PD-L1 Status Concordance – 28-8, 22C3, E1L3N, SP142, and SP263 LDTs (critical outcome [B])							
1 RCS	Very Serious	Not serious [D]	Not serious	Not serious	None	Low	
Interobserver Agreement (critical outcome [B])							
2 RCS	Very Serious	Not serious	Not serious	Not serious	None	Low	
STATEMENT 5							
Immunotherapy Response Rates by TPS (critical outcome [B])							Very Low
2 RCT, 1 post-hoc, 2 RCS	Very Serious	Not Serious	Not Serious	Not Serious	None	Low	
Immunotherapy Survival Rates by TPS (critical outcome [B])							
2 RCT, 1 post-hoc, 2 RCS	Very Serious	Serious	Not Serious	Not Serious	None	Very Low	
Interobserver Agreement – Multiple IHC Clones Stratified by TPS (critical outcome [B])							
5 RCS	Very Serious	Serious	Not serious	Not serious	None	Very Low	
STATEMENT 6							
Survival, all immunotherapy agents (critical outcome [B])							Very Low
8 RCS	Very Serious	Serious	Not serious	Not serious	None	Very Low	
Response Rates, all immunotherapy agents (critical outcome [B])							
2 RCS	Extremely serious	Not serious	Not serious	Not serious	None	Very Low	

Abbreviations: EBUS-TBNA, Endobronchial Ultrasound-guided Transbronchial needle aspiration; FNA, fine needle aspirate; IHC, immunohistochemistry; LDT, laboratory developed test; MA, meta-analysis; PCS, prospective cohort study; PD-L1, programmed death ligand-1, RCS, retrospective cohort study; RCT, randomized controlled trial; SR, systematic review; TPS, tumor proportion score.

Footnotes

A. Other category includes assessment for detection of publication bias, large effect, and confounding.
B. Outcomes were rated *a priori* as critical or important for decision making.
C. There is some inconsistency in the data, but this is likely due to differences in immunotherapy agents, PD-L1 expression cut-offs, and patient populations. Evidence was not downgraded.
D. As there is only one study included here, the assessment of inconsistency across included studies is limited.
E. There is inconsistency in the data but this is believed to be due to differences in clones and evidence was not downgraded.
F. There is inconsistency in the expression status data but the inconsistency is likely due to specimen differences and evidence was not downgraded.
G. This outcome indirectly informed the recommendation statement.

Supplemental Figure 1: Database Search Strings

Ovid, MEDLINE:

(((large cell carcinoma/ or carcinoma, non-small cell/ or exp adenocarcinoma of lung/ or (adenocarcinoma and lung).tw,kf. or (NSCLC or non-small-cell lung cancer or non-small-cell lung carcinoma).tw,kf. or (((lung or respiratory or pulmonary) and (neoplasm* or tumo?r* or cancer* or carcinoma*)) adj 3 (squamous cell or large cell or adenosquamous or Sarcomatoid)).tw. or (((lung or respiratory or pulmonary) and (neoplasm* or tumo?r* or cancer* or carcinoma*)) adj 3 (squamous cell or large cell or adenosquamous or Sarcomatoid)).kf.) **AND** (B7-H1 antigen/ or programmed cell death 1 receptor/ or programmed cell death 1 ligand.tw,kf. or (B7 H1 or PD L1 or B7H1 or B7-H1 or PDL1 or PD-L1 or PDL-1 or CD274 or CD 274).tw,kf. or (tmb or tumo?r mutation* load or tumo?r mutation* burden).tw,kf. or (mutation*adj3 (load or burden)).tw or (mutation*adj3 (load or burden)).kf. or (programmed cell death 1 or programmed cell death 1 receptor or PD1 or PD 1 or CD279 or CD 279).tw,kf.)) **NOT** ((comment or editorial/ or letter/ or case reports/ or review) or (exp animals/ not humans))) **Limit to** (English language and yr="2010-Current")

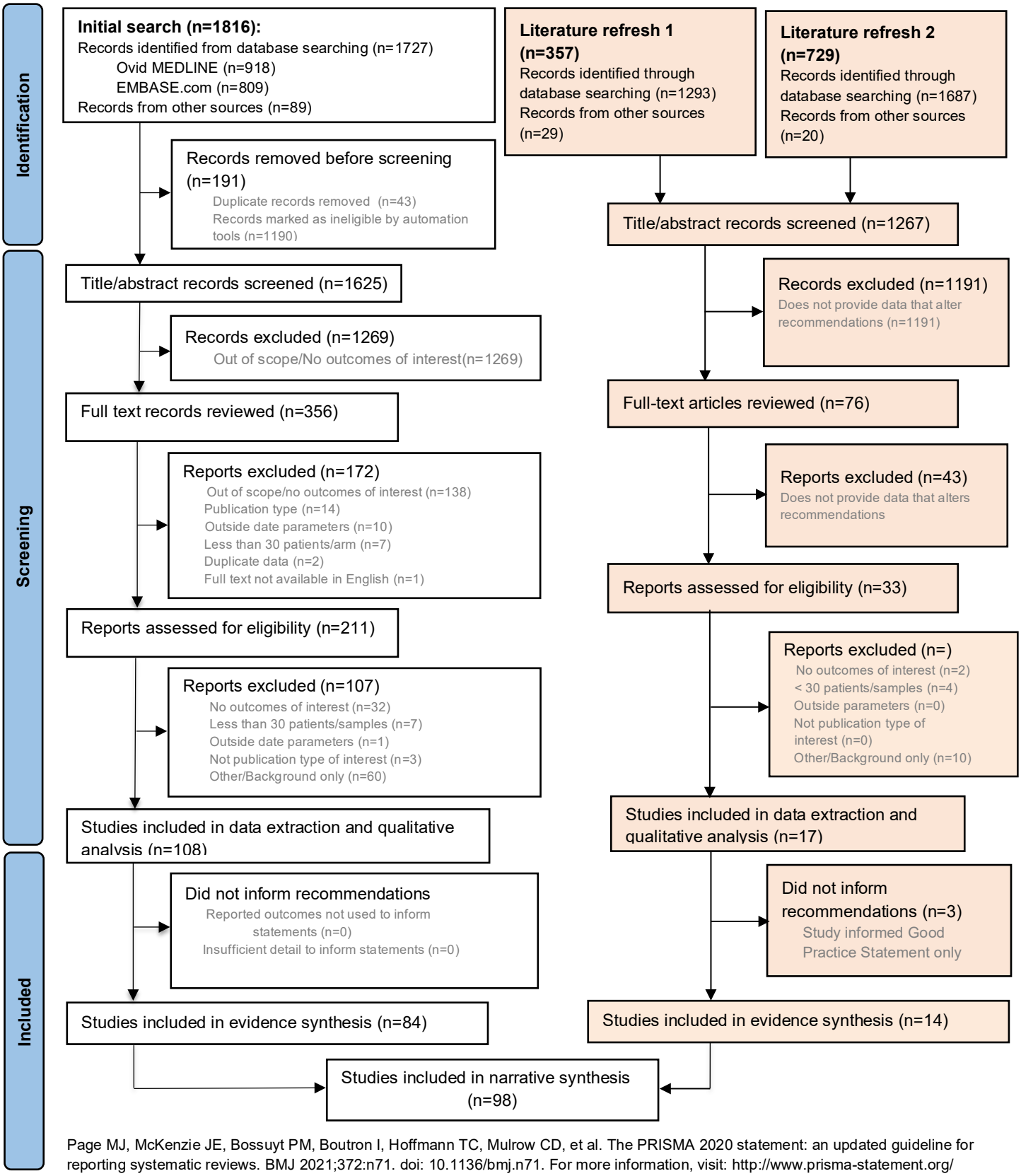
Embase:

(((('lung adenocarcinoma':ti,ab,kw OR 'large cell carcinoma':ti,ab,kw OR nsclc:ti,ab,kw OR 'non small cell lung cancer':ti,ab,kw OR 'non small cell lung carcinoma':ti,ab,kw OR (adenocarcinoma:ti,ab,kw AND (lung:ti,ab,kw OR respiratory:ti,ab,kw OR pulmonary:ti,ab,kw)) OR (squamous:ti,ab,kw AND (neoplasm*:ti,ab,kw OR tumor:ti,ab,kw OR cancer:ti,ab,kw OR carcinoma*:ti,ab,kw) AND (lung:ti,ab,kw OR respiratory:ti,ab,kw OR pulmonary:ti,ab,kw)) OR ('large cell':ti,ab,kw AND (neoplasm*:ti,ab,kw OR tumor:ti,ab,kw OR cancer:ti,ab,kw OR carcinoma*:ti,ab,kw) AND (lung:ti,ab,kw OR respiratory:ti,ab,kw OR pulmonary:ti,ab,kw)) OR (adenosquamous:ti,ab,kw AND (neoplasm*:ti,ab,kw OR tumor:ti,ab,kw OR cancer:ti,ab,kw OR carcinoma*:ti,ab,kw) AND (lung:ti,ab,kw OR respiratory:ti,ab,kw OR pulmonary:ti,ab,kw)) OR (sarcomatoid:ti,ab,kw AND (neoplasm*:ti,ab,kw OR tumor:ti,ab,kw OR cancer:ti,ab,kw OR carcinoma*:ti,ab,kw) AND (lung:ti,ab,kw OR respiratory:ti,ab,kw OR pulmonary:ti,ab,kw)) OR (pleomorphic:ti,ab,kw AND (neoplasm*:ti,ab,kw OR tumor:ti,ab,kw OR cancer:ti,ab,kw OR carcinoma*:ti,ab,kw) AND (lung:ti,ab,kw OR respiratory:ti,ab,kw OR pulmonary:ti,ab,kw))) **AND** ('programmed death 1 ligand 1':ti,ab,kw OR 'programmed death 1 receptor':ti,ab,kw OR b7h1:ti,ab,kw OR 'b7h 1':ti,ab,kw OR 'b7 h1':ti,ab,kw OR 'programmed cell death 1 ligand':ti,ab,kw OR 'tumor mutational load':ti,ab,kw OR 'tumor mutation burden':ti,ab,kw OR 'tumor mutational burden':ti,ab,kw OR 'programmed cell death 1 receptor':ti,ab,kw OR 'programmed cell death 1':ti,ab,kw OR pd1:ti,ab,kw OR 'pd 1':ti,ab,kw OR cd279:ti,ab,kw OR 'cd 279':ti,ab,kw OR pdl1:ti,ab,kw OR 'pdl 1':ti,ab,kw OR 'pd l1':ti,ab,kw OR cd274:ti,ab,kw OR 'cd 274':ti,ab,kw)) **AND** [adult]/lim **AND** 'human'/de **AND** (2021:py OR 2022:py)) **NOT** ('article'/it OR 'article in press'/it OR 'conference paper'/it)

TRIP Database:

((NSCLC) OR ("non-small-cell-lung-cancer") OR ("non small cell lung cancer") OR ("adenocarcinoma of lung") OR("large cell carcinoma")) AND (("B7-H1 antigen") OR ("programmed cell death ligand") OR (B7H1) OR (B7-H1) OR (CD279) OR (CD-279) OR (TMB) OR ("tumor mutation burden") OR ("tumor mutation load") OR ("tumor mutational load") OR ("tumor mutational burden") OR ("programmed cell death receptor") OR ("programmed cell death") OR (PD1) OR (CD274) OR (CD-274)) from:2010 to:2019

Supplemental Figure 2: Literature Review Flow Diagram



Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71. For more information, visit: http://www.prisma-statement.org/

Glossary of Terms

Acceptability—Acceptability reflects who benefits (or is harmed) and who pays (or saves); and when the benefits, adverse effects, and costs occur (and the discount rates of key stakeholders, eg, politicians may have a high discount rate for anything that occurs beyond the next election). For the Evidence to decision (EtD) framework, the expert panel considered target users of the guideline. The less acceptable an option is to key stakeholders, the less likely it is that it should be recommended, or if it is recommended, the more likely it is that the recommendation should include an implementation strategy to address concerns about acceptability.

Accuracy—The degree of correctness or true values of a given laboratory result comparing to a gold standard. Accuracy also implies freedom from error.

Advanced stage—Includes patients with stage IIIB or IV disease, generally considered to include locally advanced/unresectable and metastatic cancer, respectively.

Advisory Panel—Group established to provide additional expertise needed outside of the expert panel. Their primary role is to review the draft guideline during key stages of development however they do not hold any formal decision-making capabilities or have voting rights. Advisory Panel members generally do not author the guideline; however, these decisions may be made on a case-by-case basis, as determined by the primary authors, subject to the conflict of interest (COI) disclosures and policies of the publishing journals. Advisory Panel membership may include individuals with professional expertise from other vested organizations including but not limited to a patient advocate among others.

ALK—anaplastic lymphoma receptor tyrosine kinase

AMSTAR (Assessing the Methodological Quality of Systematic Reviews)—A validated quality assessment tool for systematic reviews.

B7H1—B7 Homolog 1 also known as PD-L1

Benefit—A valued or desired outcome. In EtD, the expert panel considers both the magnitude of the benefits as well as the importance of that benefit to both clinicians and patients.

BRAF—B-Raf Proto-Oncogene

CD274—gene that encodes PD-L1

Companion Diagnostics (CDx)—According to the FDA, a “companion diagnostic is a medical device, often an in vitro diagnostic (IVD), which provides information that is essential for the safe and effective use of a corresponding drug or biological product”.

Combined Positive Score—The number of positive tumor cells, lymphocytes and macrophages, divided by the total number of viable tumor cells multiplied by 100.

Concordance—The degree of agreement between two quantitative methods or assays.

Confidence Interval (CI)—The 95% confidence interval is a range of values that we can be 95% certain contains the point statistic.

Conflict of Interest (COI)—A divergence between an individual’s private interests and his or her professional obligations such that an independent observer might reasonably question whether the individual’s professional actions or decisions are motivated by personal gain, such as financial, academic advancement, clinical revenue streams, or community standing. This includes financial and intellectual relationships that may impact an individual’s approach a scientific question with an open mind.

Cost—In this guideline, the discussion on cost pertains to the use of resources for an intervention or a recommendation.

Disease-free survival—The measure of time after treatment during which no sign of disease is found.

Early stage—Includes patients with stage I disease, generally considered to include tumors that has not spread to the lymph nodes. These patients have tumors that are routinely resected.

EGFR— epidermal growth factor receptor

Equity—Health equity is the attainment of the highest level of health for all people. For the EtD, the EP deliberated any advantages or disadvantages for any group or setting in relation to the recommendation being considered. The EP considered any differences in baseline conditions across groups or settings that affect the absolute effectiveness of the recommendation or the importance of the problem for disadvantaged groups or settings. The EP discussed any important considerations that should be made when implementing the recommendations to ensure that inequities are reduced or eliminated.

Evidence-to-decision framework (EtD)—The purpose of this framework is to help panels developing guidelines move from evidence to recommendations. It is intended to inform panel members' judgements about the pros and cons of each intervention that is considered; ensure that important factors that determine a recommendation are considered; provide a concise summary of the best available research evidence to inform judgements about each criterion; help structure discussion and identify reasons for disagreements; and make the basis for recommendations transparent to guideline users[1]

Expert Panel (EP)—Group established to approve key questions as defined by the guideline co-chairs, assist in the systematic review of the evidence, develop the draft recommendations, and write the final recommendations including full manuscript. The expert panel is overseen by the co-chairs, holds authorship attribution on the final guideline manuscript and is usually comprised of multidisciplinary topic experts. Expert Panel membership may include individuals with professional expertise from other vested organizations.

Feasibility—The capability of an intervention or an action to be accomplished or implemented. The less feasible an option is, the less likely it is that it should be recommended. For the EtD, the EP considered barriers that are likely to limit the feasibility of implementing the recommendation.

Grading of Recommendations Assessment, Development and Evaluation (GRADE)—An internationally accepted and validated approach to grading quality of evidence and strength or recommendations.

Harms—A risk or injury occurring as a result of an intervention. In EtD, the expert panel considers both the magnitude of the harms as well as the importance of that harm to both clinicians and patients.

Imprecision—A domain of the GRADE strength of evidence assessment. Imprecision results when evidence carries a wide confidence interval around the estimate of effect.

Inconsistency—A domain of the GRADE strength of evidence assessment. Inconsistency refers to an unexplained heterogeneity of results across studies informing a guidance statement.

Indirectness—A domain of the GRADE strength of evidence assessment. Indirectness refers to evidence that does not directly inform the PICO elements.

Interobserver Agreement—The degree to which two or more independent observers report the same values after measuring the same events.

Intraobserver Agreement—The degree to which two or more values are reported after being measured by the same observer.

KRAS— Kirsten rat sarcoma viral oncogene homolog protein

Locally-advanced NSCLC—Defined by the National Institutes of Health as stage III subclassification into stages IIIA, IIIB, and IIIC.

Meta-Analysis (MA) —Statistical procedure for combining data from multiple studies. Outcomes from a meta-analysis may include a more precise estimate of the effect of treatment or risk factor for disease, or other outcomes, than any individual study contributing to the pooled analysis.

Negative Predictive Value (NPV) —The predictive value of a negative result. This value corresponds to the percentage of true negative patients among those given a negative test result.

Overall Survival (OS)—The length of time from either the date of diagnosis or the start of treatment to death from any cause.

Outcomes—Outcomes are the potential benefits or harms. Outcomes that are considered to be important to those affected by the intervention, and which are important to making a recommendation or decision. Consultation with those affected by an intervention (such as patients and their caretakers) or other members of the public may be used to select the important outcomes. A review of the literature may also be carried out to inform the selection of the important outcomes. The importance (or value) of each outcome in relation to the other outcomes should also be considered. This is the relative importance of the outcome.

PICO—A validated approach to developing guideline research questions that frames the population of interest (P), interventions (I) under consideration, possible comparisons (C), and relevant research outcomes (O).

Problem—In the EtD framework, the EP considered the priority of the problem a recommendation is addressing. The EP considered if the consequences of the problem are serious and if addressing the problem is urgent. Serious problems are more likely that an option which addresses the problem should be a priority (e.g., diseases that are fatal or disabling are likely to be a higher priority than diseases that only cause minor distress). The more people who are affected, the more likely it is that an option that addresses the problem should be a priority.

Progression-Free Survival (PFS)—The length of time from treatment to disease progression or death.

Prospective Cohort Study (PCS) —Study design that enrolls a cohort of subjects and watches those subjects over a time period. A prospective study watches for outcomes during the study period and relates those outcomes to prior exposure or clinical characteristic.

Positive Predictive Value (PPV)—The predictive value of a positive result. This value corresponds to the percentage of true positive patients among those given a positive test result.

Randomized Controlled Trial (RCT)—Study design that randomly assigns subjects into an experimental group or a control group. Subjects are followed to determine effectiveness of the experimental intervention with outcomes measured at specific time-points.

Recurrence-Free Survival (RFS)—The length of time from treatment to disease recurrence or death.

Resectable NSCLC—Patients with resectable disease in stages I-IIIB have surgery as the primary treatment option.

Retrospective Cohort Study (RCS)—Study design that enrolls a cohort of subjects based on a known outcome and looks backwards to correlate prior exposure or clinical characteristic to that outcome.

Risk of Bias—The risk of systematic error or deviation from the truth within a scientific study.

ROBINS-I (Risk of Bias In Non-Randomized Studies – of Intervention)—A validated quality assessment tool for observational studies.

ROS1—ROS Proto-Oncogene 1

Sensitivity—The probability that a diagnostic test identifies patients who are in fact positive for a disease. The value corresponds to the percentage of true positive results demonstrated by an assay among those who are truly positive.

Specificity—The probability that a diagnostic test identifies patients who are in fact negative for a disease. The value corresponds to the percentage of true negative results demonstrated by an assay among those who are truly negative.

Systematic Review (SR) —A systematic review summarizes the results of available carefully designed healthcare studies and provides a high level of evidence on the effectiveness of healthcare interventions. Judgments may be made about the evidence and inform recommendations for healthcare.

Time to Recurrence—The length of time from treatment to disease recurrence.

Tumor Proportion Score (TPS)—The percentage of viable tumor cells showing partial or complete membrane staining at any intensity.

References

1. College of American Pathologists Pathology and Laboratory Quality Center for Evidence-based Guidelines. Evidence-based guideline development methodology manual. College of American Pathologists website. Published August 2020. Accessed November 7, 2023. https://documents.cap.org/documents/cap-center-ebg-developmentmanual.pdf
2. Bramer WM, Giustini D, de Jonge GB, Holland L, Bekhuis T. De-duplication of database search results for systematic reviews in EndNote. *J Med Libr Assoc*. 2016;104(3):240-243. doi:10.3163/1536-5050.104.3.014.
3. Schuenemann H BJ, Guyatt G, Oxman A. Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach; 2013. https://gdt.gradepro.org/app/handbook/handbook.html
4. Alonso-Coello P, Schünemann HJ, Moberg J, et al. GRADE Evidence to Decision (EtD) frameworks: A systematic and transparent approach to making well informed healthcare choices. 1: Introduction. *BMJ*. 2016;353:i2016. doi:10.1136/bmj.i2016.
5. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*. 2017;358:j4008. doi:10.1136/bmj.j4008.
6. Higgins JP, Altman DG, Gøtzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011;343:d5928. doi:10.1136/bmj.d5928.
7. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. 2016;355:i4919. doi:10.1136/bmj.i4919.
8. Cao R, Ma J-T, Zhang S-L, et al. Rational application of the first-line chemotherapy and immune checkpoint inhibitors in advanced nonsmall cell lung cancer: A meta-analysis. *Cancer Med*. 2019;8(11):5033-5046. doi:10.1002/cam4.2407
9. Kim R, Keam B, Hahn S, et al. First-line Ppmbrolizumab versus pembrolizumab plus chemotherapy versus chemotherapy alone in non-small-cell lung cancer: A systematic review and network meta-analysis. *Clin Lung Cancer*. 2019;20(5):331-338.e4. doi:10.1016/j.cllc.2019.05.009
10. Borghaei H, Paz-Ares L, Horn L, et al. Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. *N Engl J Med*. 2015;373(17):1627-1639. doi:10.1056/NEJMoa1507643
11. Brahmer J, Reckamp KL, Baas P, et al. Nivolumab versus docetaxel in advanced squamous-cell non-small-cell lung cancer. *N Engl J Med*. 2015;373(2):123-135. doi:10.1056/NEJMoa1504627
12. Fehrenbacher L, Spira A, Ballinger M, et al. Atezolizumab versus docetaxel for patients with previously treated non-small-cell lung cancer (POPLAR): A multicentre, open-label, phase 2 randomised controlled trial. *Lancet*. 2016;387(10030):1837-1846. doi:10.1016/S0140-6736(16)00587-0
13. Garon EB, Hellmann MD, Rizvi NA, et al. Five-year overall survival for patients with advanced non–small-ell lung cancer treated with pembrolizumab: Results from the phase I KEYNOTE-001 study. *J Clin Oncol*. 2019;37(28):2518-2527. doi:10.1200/JCO.19.00934
14. Hellmann MD, Paz-Ares L, Bernabe Caro R, et al. Nivolumab plus ipilimumab in advanced non-small-cell lung cancer. *N Engl J Med*. 2019;381(21):2020-2031. doi:10.1056/NEJMoa1910231
15. Mok TSK, Wu YL, Kudaba I, et al. Pembrolizumab versus chemotherapy for previously untreated, PD-L1-expressing, locally advanced or metastatic non-small-cell lung cancer (KEYNOTE-042): A randomised, open-label, controlled, phase 3 trial. *Lancet*. 2019;393(10183):1819-1830. doi:10.1016/S0140-6736(18)32409-7
16. Ready N, Hellmann MD, Awad MM, et al. First-line nivolumab plus ipilimumab in advanced non-small-cell lung cancer (CheckMate 568): Outcomes by programmed death ligand 1 and tumor mutational burden as biomarkers. *J Clin Oncol*. 2019;37(12):992-1000. doi:10.1200/JCO.18.01042
17. Rittmeyer A, Barlesi F, Waterkamp D, et al. Atezolizumab versus docetaxel in patients with previously treated non-small-cell lung cancer (OAK): A phase 3, open-label, multicentre randomised controlled trial. *Lancet*. 2017;389(10066):255-265. doi:10.1016/S0140-6736(16)32517-X
18. West H, McCleod M, Hussein M, et al. Atezolizumab in combination with carboplatin plus nab-paclitaxel chemotherapy compared with chemotherapy alone as first-line treatment for metastatic non-squamous non-small-cell lung cancer (IMpower130): A multicentre, randomised, open-label, phase 3 trial. *Lancet Oncol*. 2019;20(7):924-937. doi:10.1016/S1470-2045(19)30167-6

19. Reck M, Rodriguez-Abreu D, Robinson AG, et al. Five-year outcomes with pembrolizumab versus chemotherapy for metastatic non-small-cell lung cancer with PD-L1 tumor proportion score >= 50. *J Clin Oncol*. 2021;39(21):2339-2349. doi:10.1200/JCO.21.00174

20. Paz-Ares LG, Ramalingam SS, Ciuleanu TE, et al. First-line nivolumab plus ipilimumab in advanced NSCLC: 4-year outcomes from the randomized, open-label, phase 3 CheckMate 227 Part 1 trial. *J Thorac Oncol:* 2022;17(2):289-308. doi:10.1016/j.jtho.2021.09.010

21. Park K, Ozguroglu M, Vansteenkiste J, et al. Avelumab versus docetaxel in patients with platinum-treated advanced NSCLC: 2-year follow-up from the JAVELIN Lung 200 phase 3 trial. *J Thorac Oncol:* 2021;16(8):1369-1378. doi:10.1016/j.jtho.2021.03.009

22. Jassem J, de Marinis F, Giaccone G, et al. Updated overall survival analysis from IMpower110: Atezolizumab versus platinum-based chemotherapy in treatment-naive programmed death-ligand 1-selected NSCLC. *J Thorac Oncol*. 2021;16(11):1872-1882. doi:10.1016/j.jtho.2021.06.019

23. Herbst RS, Garon EB, Kim DW, et al. Five year survival update from KEYNOTE-010: Pembrolizumab versus docetaxel for previously treated, programmed death-ligand 1-positive advanced NSCLC. *J Thorac Oncol.* 2021;16(10):1718-1732. doi:10.1016/j.jtho.2021.05.001

24. Garon EB, Rizvi NA, Hui R, et al. Pembrolizumab for the treatment of non-small-cell lung cancer. *N Engl J Med*. 2015;372(21):2018-2028. doi:10.1056/NEJMoa1501824

25. Herbst RS, Baas P, Perez-Gracia JL, et al. Use of archival versus newly collected tumor samples for assessing PD-L1 expression and overall survival: An updated analysis of KEYNOTE-010 trial. *Ann Oncol*. 2019;30(2):281-289. doi:10.1093/annonc/mdy545

26. Horn L, Gettinger SN, Gordon MS, et al. Safety and clinical activity of atezolizumab monotherapy in metastatic non-small-cell lung cancer: final results from a phase I study. *Eur J Cancer*. 2018;101:201-209. doi:10.1016/j.ejca.2018.06.031

27. Hui R, Garon EB, Goldman JW, et al. Pembrolizumab as first-line therapy for patients with PD-L1-positive advanced non-small cell lung cancer: A phase 1 trial. *Ann Oncol*. 2017;28(4):874-881. doi:10.1093/annonc/mdx008

28. Antonia SJ, Balmanoukian A, Brahmer J, et al. Clinical activity, tolerability, and long-term follow-up of durvalumab in patients with advanced NSCLC. *J Thorac Oncol*. 2019;14(10):1794-1806. doi:10.1016/j.jtho.2019.06.010

29. Gettinger S, Horn L, Jackman D, et al. Five-year follow-up of nivolumab in previously treated advanced non–small-cell lung cancer: Results from the CA209-003 study. *J Clin Oncol.* 2018;36(17):1675-1684. doi:10.1200/JCO.2017.77.0412

30. Gettinger S, Rizvi NA, Chow LQ, et al. Nivolumab monotherapy for first-line treatment of advanced non-small-cell lung cancer. *J Clin Oncol*. 2016;34(25):2980-2987. doi:10.1200/JCO.2016.66.9929

31. Peters S, Gettinger S, Johnson ML, et al. Phase II trial of atezolizumab as first-line or subsequent therapy for patients with programmed death-ligand 1-selected advanced non-small-cell lung cancer (BIRCH). *J Clin Oncol*. 2017;35(24):2781-2789. doi:10.1200/JCO.2016.71.9476

32. Aguilar EJ, Ricciuti B, Gainor JF, et al. Outcomes to first-line pembrolizumab in patients with non-small cell lung cancer and very high PD-L1 expression. *Ann Oncol*. 2019;30(10):1653-1659. doi:10.1093/annonc/mdz288

33. Ahn BC, Pyo KH, Xin CF, et al. Comprehensive analysis of the characteristics and treatment outcomes of patients with non-small cell lung cancer treated with anti-PD-1 therapy in real-world practice. *J Cancer Res Clin Oncol*. 2019;145(6):1613-1623. doi:10.1007/s00432-019-02899-y

34. Edahiro R, Kanazu M, Kurebe H, et al. Clinical outcomes in non-small cell lung cancer patients with an ultra-high expression of programmed death ligand-1 treated using pembrolizumab as a first-line therapy: A retrospective multicenter cohort study in Japan. *PLoS One*. 2019;14(7):e0220570. doi:10.1371/journal.pone.0220570

35. Kim HS, Cha H, Kim J, et al. Genomic scoring to determine clinical benefit of immunotherapy by targeted sequencing. *Eur J Cancer*. 2019;120:65-74. doi:10.1016/j.ejca.2019.08.001

36. Lin SY, Yang CY, rLiao BC, et al. Tumor PD-L1 expression and clinical outcomes in advanced-stage non-small cell lung cancer patients treated with nivolumab or pembrolizumab: Real-world data in Taiwan. *J Cancer*. 2018;9(10):1813-1820. doi:10.7150/jca.24985

37. Oya Y, Yoshida T, Kuroda H, et al. Predictive clinical parameters for the response of nivolumab in pretreated advanced non-small-cell lung cancer. *Oncotarget*. 2017;8(61):103117-103128. doi:10.18632/oncotarget.21602

38. Rizvi H, Sanchez-Vega F, La K, et al. Molecular determinants of response to anti-programmed cell death (PD)-1 and anti-programmed death-ligand 1 (PD-L1) blockade in patients with non-small-cell lung

cancer profiled with targeted next-generation sequencing. *J Clin Oncol*. 2018;36(7):633-641. doi:10.1200/JCO.2017.75.3384

39. Tamiya M, Tamiya A, Hosoya K, et al. Efficacy and safety of pembrolizumab as first-line therapy in advanced non-small cell lung cancer with at least 50% PD-L1 positivity: A multicenter retrospective cohort study (HOPE-001). *Invest New Drugs*. 2019;37(6):1266-1273. doi:10.1007/s10637-019-00843-y

40. Fujimoto D, Sato Y, Uehara K, et al. Predictive performance of four programmed cell death ligand 1 assay systems on nivolumab response in previously treated patients with non-small cell lung cancer. *J Thorac Oncol*. 2018;13(3):377-386. doi:10.1016/j.jtho.2017.11.123

41. Tseng J-S, Yang T-Y, Wu C-Y, et al. Characteristics and predictive value of PD-L1 status in real-world non-small cell lung cancer patients. *J Immunother*. 2018;41(6):292-299. doi:10.1097/CJI.0000000000000226

42. Barlesi F, Vansteenkiste J, Spigel D, et al. Avelumab versus docetaxel in patients with platinum-treated advanced non-small-cell lung cancer (JAVELIN Lung 200): An open-label, randomised, phase 3 study. *Lancet Oncol*. 2018;19(11):1468-1479. doi:10.1016/S1470-2045(18)30673-9

43. Herbst RS, Baas P, Kim DW, et al. Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): A randomised controlled trial. *Lancet*. 2016;387(10027):1540-1550. doi:10.1016/S0140-6736(15)01281-7

44. Herbst RS, Giaccone G, de Marinis F, et al. Atezolizumab for first-line treatment of PD-L1-selected patients with NSCLC. *N Engl J Med*. 2020;383(14):1328-1339. doi:10.1056/NEJMoa1917346

45. Fitzgibbons PL, Bradley LA, Fatheree LA, et al. Principles of analytic validation of immunohistochemical assays: Guideline from the College of American Pathologists Pathology and Laboratory Quality Center. *Arch Pathol Lab Med*. 2014;138(11):1432-1443. doi:10.5858/arpa.2013-0610-CP

46. Vigliar E, Malapelle U, Iaccarino A, et al. PD-L1 expression on routine samples of non-small cell lung cancer: Results and critical issues from a 1-year experience of a centralised laboratory. *J Clin Pathol*. 2019;72(6):412-417. doi:10.1136/jclinpath-2019-205732

47. Wang G, Ionescu DN, Lee CH, et al. PD-L1 testing on the EBUS-FNA cytology specimens of non-small cell lung cancer. *Lung Cancer*. 2019;136:1-5. doi:10.1016/j.lungcan.2019.07.033

48. Chan AWH, Tong JHM, Kwan JSH, et al. Assessment of programmed cell death ligand-1 expression by 4 diagnostic assays and its clinicopathological correlation in a large cohort of surgical resected non-small cell lung carcinoma. *Mod Pathol*. 2018;31(9):1381-1390. doi:10.1038/s41379-018-0053-3

49. Elfving H, Mattsson JSM, Lindskog C, Backman M, Menzel U, Micke P. Programmed cell death ligand 1 immunohistochemistry: A concordance study between surgical specimen, biopsy, and tissue microarray. *Clin Lung Cancer*. 2019;20(4):258-262.e1. doi:10.1016/j.cllc.2019.02.012

50. Gradecki SE, Grange JS, Stelow EB. Concordance of PD-L1 expression between core biopsy and resection specimens of non-small cell lung cancer. *Am J Surg Pathol*. 2018;42(8):1090-1094. doi:10.1097/PAS.0000000000001085

51. Grosu HB, Arriola A, Stewart J, et al. PD-L1 detection in histology specimens and matched pleural fluid cell blocks of patients with NSCLC. *Respirology*. 2019;24(12):1198-1203. doi:10.1111/resp.13614

52. Hernandez A, Brandler TC, Zhou F, Moreira AL, Schatz-Siemers N, Simsir A. Assessment of programmed death–ligand 1 (PD-L1) immunohistochemical expression on cytology specimens in non–small cell lung carcinoma: A comparative study with paired surgical specimens. *Am J Clin Pathol*. 2019;151(4):403-415. doi:10.1093/ajcp/aqy164

53. Ilie M, Long-Mira E, Bence C, et al. Comparative study of the PD-L1 status between surgically resected specimens and matched biopsies of NSCLC patients reveal major discordances: A potential issue for anti-PD-L1 therapeutic strategies. *Ann Oncol*. 2016;27(1):147-153. doi:10.1093/annonc/mdv489

54. Keller MD, Neppl C, Irmak Y, et al. Adverse prognostic value of PD-L1 expression in primary resected pulmonary squamous cell carcinomas and paired mediastinal lymph node metastases. *Mod Pathol*. 2018;31(1):101-110. doi:10.1038/modpathol.2017.111

55. Kim HR, Cha YJ, Hong MH, et al. Concordance of programmed death-ligand 1 expression between primary and metastatic non-small cell lung cancer by immunohistochemistry and RNA in situ hybridization. *Oncotarget*. 2017;8(50):87234-87243. doi:10.18632/oncotarget.20254

56. Kim S, Koh J, Kwon D, et al. Comparative analysis of PD-L1 expression between primary and metastatic pulmonary adenocarcinomas. *Eur J Cancer*. 2017;75:141-149. doi:10.1016/j.ejca.2017.01.004

57. Kuempers C, van der Linde LIS, Reischl M, et al. Comparison of PD-L1 expression between paired cytologic and histologic specimens from non-small cell lung cancer patients. *Virchows Archiv*. 2020;476(2):261-271. doi:10.1007/s00428-019-02632-7

58. Mei P, Shilo K, Wei L, Shen R, Tonkovich D, Li Z. Programmed cell death ligand 1 expression in cytologic and surgical non–small cell lung carcinoma specimens from a single institution: Association with clinicopathologic features and molecular alterations. *Cancer Cytopathol*. 2019;127(7):447-457. doi:10.1002/cncy.22140

59. Munari E, Zamboni G, Lunardi G, et al. PD-L1 expression heterogeneity in non–small cell lung cancer: defining criteria for harmonization between biopsy specimens and whole sections. *J Thorac Oncol*. 2018;13(8):1113-1120. doi:10.1016/j.jtho.2018.04.017

60. Munari E, Zamboni G, Sighele G, et al. Expression of programmed cell death ligand 1 in non–small cell lung cancer: Comparison between cytologic smears, core biopsies, and whole sections using the SP263 assay. *Cancer Cytopathol*. 2019;127(1):52-61. doi:10.1002/cncy.22083

61. Noll B, Wang WL, Gong Y, et al. Programmed death ligand 1 testing in non–small cell lung carcinoma cytology cell block and aspirate smear preparations. *Cancer Cytopathol*. 2018;126(5):342-352. doi:10.1002/cncy.21987

62. Rehman JA, Han G, Carvajal-Hausdorf DE, et al. Quantitative and pathologist-read comparison of the heterogeneity of programmed death-ligand 1 (PD-L1) expression in non-small cell lung cancer. *Mod Pathol*. 2017;30(3):340-349. doi:10.1038/modpathol.2016.186

63. Sakata KK, Midthun DE, Mullon JJ, et al. Comparison of programmed death ligand-1 immunohistochemical staining between endobronchial ultrasound transbronchial needle aspiration and resected lung cancer specimens. *Chest*. 2018;154(4):827-837. doi:10.1016/j.chest.2018.07.017

64. Sheffield BS, Fulton R, Kalloger SE, et al. Investigation of PD-L1 biomarker testing methods for PD-1 axis inhibition in non-squamous non–small cell lung cancer. *J Histochem Cytochem*. 2016;64(10):587-600. doi:10.1369/0022155416665338

65. Téglási V, Pipek O, Lózsa R, et al. PD-L1 expression of lung cancer cells, unlike infiltrating immune cells, is stable and unaffected by therapy during brain metastasis. *Clin Lung Cancer*. 2019;20(5):363-369.e2. doi:10.1016/j.cllc.2019.05.008

66. Uruga H, Bozkurtlar E, Huynh TG, et al. Programmed cell death ligand (PD-L1) expression in stage II and III lung adenocarcinomas and nodal metastases. *J Thorac Oncol*. 2017;12(3):458-466. doi:10.1016/j.jtho.2016.10.015

67. Ilie M, Juco J, Huang L, Hofman V, Khambata-Ford S, Hofman P. Use of the 22C3 anti-programmed death-ligand 1 antibody to determine programmed death-ligand 1 expression in cytology samples obtained from non-small cell lung cancer patients. *Cancer Cytopathol*. 2018;126(4):264-274. doi:10.1002/cncy.21977.

68. Munari E, Zamboni G, Lunardi G, et al. PD-L1 expression comparison between primary and relapsed non-small cell lung carcinoma using whole sections and clone SP263. *Oncotarget*. 2018;9(54):30465-30471. doi:10.18632/oncotarget.25770

69. Skov BG, Skov T. Paired comparison of PD-L1 expression on cytologic and histologic specimens from malignancies in the lung assessed with PD-L1 IHC 28-8pharmDx and PD-L1 IHC 22C3pharmDx. *Appl Immunohistochem Mol Morphol.* 2017;25(7):453-459. doi:10.1097/PAI.0000000000000540

70. Torous VF, Rangachari D, Gallant BP, Shea M, Costa DB, VanderLaan PA. PD-L1 testing using the clone 22C3 pharmDx kit for selection of patients with non-small cell lung cancer to receive immune checkpoint inhibitor therapy: Are cytology cell blocks a viable option? *J Am Soc Cytopathol*. 2018;7(3):133-141. doi:10.1016/j.jasc.2018.02.003

71. Wang H, Agulnik J, Kasymjanova G, et al. The metastatic site does not influence PD-L1 expression in advanced non-small cell lung carcinoma. *Lung Cancer*. 2019;132:36-38. doi:10.1016/j.lungcan.2019.04.009

72. 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(6):1417-1422. doi:10.1093/annonc/mdy126

73. Russell-Goldman E, Kravets S, Dahlberg SE, Sholl LM, Vivero M. Cytologic-histologic correlation of programmed death-ligand 1 immunohistochemistry in lung carcinomas. *Cancer Cytopathol*. 2018;126(4):253-263. doi:10.1002/cncy.21973

74. Daverio M, Patrucco F, Gavelli F, et al. Comparative analysis of programmed death ligand 1 expression in paired cytologic and histologic specimens of non-small cell lung cancer. *Cancer Cytopathol*. 2020;128(8):580-588. doi:10.1002/cncy.22292

75. Hernandez A, Brandler TC, Chen F, et al. Scoring of programmed death-ligand 1 immunohistochemistry on cytology cell block specimens in non-small cell lung carcinoma. *Am J Clin Pathol*. 2020;154(4):517-524. doi:10.1093/ajcp/aqaa073

76. Lou SK, Ko HM, Kinoshita T, et al. Implementation of PD-L1 22C3 IHC pharmDxTM in cell block preparations of lung cancer: concordance with surgical resections and technical validation of cytoLyt R prefixation. *Acta Cytol*. 2020;64(6):577-587. doi:10.1159/000508628

77. Lozano MD, Abengozar-Muela M, Echeveste JI, et al. Programmed death-ligand 1 expression on direct Pap-stained cytology smears from non-small cell lung cancer: Comparison with cell blocks and surgical resection specimens. *Cancer Cytopath*. 2019;127(7):470-480. doi:10.1002/cncy.22155

78. Song SG, Lee J, Koh J, Kim S, Chung DH, Jeon YK. Utility of PD-L1 immunocytochemistry using body-fluid cell blocks in patients with non-small-cell lung cancer. *Diagn Cytopath*. 2020;48(4):291-299. doi:10.1002/dc.24379

79. Zou Y, Xu L, Tang Q, et al. Cytology cell blocks from malignant pleural effusion are good candidates for PD-L1 detection in advanced NSCLC compared with matched histology samples. *BMC cancer*. 2020;20(1):344. doi:10.1186/s12885-020-06851-z

80. Brunnström H, Johansson A, Westbom-Fremer S, et al. PD-L1 immunohistochemistry in clinical diagnostics of lung cancer: Inter-pathologist variability is higher than assay variability. *Mod Path*. 2017;30(10):1411-1421. doi:10.1038/modpathol.2017.59

81. Cooper WA, Russell PA, Cherian M, et al. Intra- and interobserver reproducibility assessment of PD-L1 biomarker in non–small cell lung cancer. *Clin Cancer Res*. 2017;23(16):4569-4577. doi:10.1158/1078-0432.CCR-17-0151

82. Krawczyk P, Jarosz B, Kucharczyk T, et al. Immunohistochemical assays incorporating SP142 and 22C3 monoclonal antibodies for detection of PD-L1 expression in NSCLC patients with known status of EGFR and ALK genes. *Oncotarget*. 2017;8(38):64283-64293. doi:10.18632/oncotarget.19724

83. Pang C, Yin L, Zhou X, et al. Assessment of programmed cell death ligand-1 expression with multiple immunohistochemistry antibody clones in non-small cell lung cancer. *J Thorac Dis*. 2018;10(2):816-824. doi:10.21037/jtd.2018.01.124

84. Saito T, Tsuta K, Ishida M, et al. Comparative study of programmed cell death ligand-1 immunohistochemistry assays using 22C3 and 28-8 antibodies for non-small cell lung cancer: Analysis of 420 surgical specimens from Japanese patients. *Lung Cancer*. 2018;125:230-237. doi:10.1016/j.lungcan.2018.10.005

85. Song P, Guo L, Li W, Zhang F, Ying J, Gao S. Clinicopathologic correlation with expression of PD-L1 on both tumor cells and tumor-infiltrating immune cells in patients with non-small cell lung cancer. *J Immunother*. 2019;42(1):23-28. doi:10.1097/CJI.0000000000000249

86. Yeo MK, Choi SY, Seong IO, Suh KS, Kim JM, Kim KH. Association of PD-L1 expression and PD-L1 gene polymorphism with poor prognosis in lung adenocarcinoma and squamous cell carcinoma. *Human Path*. 2017;68:103-111. doi:10.1016/j.humpath.2017.08.016

87. Beck KS, Kim SJ, Kang JH, Han DH, Jung JI, Lee KY. CT-guided transthoracic needle biopsy for evaluation of PD-L1 expression: Comparison of 22C3 and SP263 assays. *Thorac Cancer*. 2019;10(7):1612-1618. doi:10.1111/1759-7714.13126

88. Fujimoto D, Yamashita D, Fukuoka J, et al. Comparison of PD-L1 assays in non-small cell lung cancer: 22C3 pharmDx and SP263. *Anticancer Res*. 2018;38(12):6891-6895. doi:10.21873/anticanres.13065

89. Humphries MP, McQuaid S, Craig SG, et al. Critical appraisal of programmed death ligand 1 reflex diagnostic testing: Current standards and future opportunities. *J Thorac Oncol*. 2019;14(1):45-53. doi:10.1016/j.jtho.2018.09.025

90. Marchetti A, Barberis M, Franco R, et al. Multicenter comparison of 22C3 PharmDx (Agilent) and SP263 (Ventana) assays to test PD-L1 expression for NSCLC patients to be treated with immune Checkpoint Inhibitors. *J Thorac Oncol*. 2017;12(11):1654-1663. doi:10.1016/j.jtho.2017.07.031

91. Munari E, Zamboni G, Lunardi G, et al. PD-L1 expression in non-small cell lung cancer: Evaluation of the diagnostic accuracy of a laboratory-developed test using clone E1L3N in comparison with 22C3 and SP263 assays. *Human Path*. 2019;90:54-59. doi:10.1016/j.humpath.2019.05.003

92. Park HY, Oh IJ, Kho BG, et al. Clinical characteristics of Korean patients with lung cancer who have programmed death-ligand 1 expression. *Tuberculosis Resp Dis (Seoul)*. 2019;82(3):227-233. doi:10.4046/trd.2018.0070

93. Parra ER, Villalobos P, Mino B, Rodriguez-Canales J. Comparison of different antibody clones for immunohistochemistry detection of programmed cell death ligand 1 (PD-L1) on non-small cell lung Carcinoma. *Appl Immunohistochem Mol Morphol* 2018;26(2):83-93. doi:10.1097/PAI.0000000000000531

94. Ratcliffe MJ, Sharpe A, Midha A, et al. Agreement between programmed cell death ligand-1 diagnostic assays across multiple protein expression cutoffs in non-small cell lung cancer. *Clin Cancer Res*. 2017;23(14):3585-3591. doi:10.1158/1078-0432.CCR-16-2375

95. Rimm DL, Han G, Taube JM, et al. A prospective, multi-institutional, pathologist-based assessment of 4 immunohistochemistry assays for PD-L1 expression in non-small cell lung cancer. *JAMA Oncol*. 2017;3(8):1051-1058. doi:10.1001/jamaoncol.2017.0013

96. Xu H, Lin G, Huang C, et al. Assessment of concordance between 22C3 and SP142 immunohistochemistry assays regarding PD-L1 expression in non-small cell lung cancer. *Sci Rep*. 2017;7(1):16956. doi:10.1038/s41598-017-17034-5

97. Tsao MS, Kerr KM, Kockx M, et al. PD-L1 immunohistochemistry comparability study in real-life clinical samples: Results of Blueprint phase 2 project. *J Thorac Oncol*. 2018;13(9):1302-1311. doi:10.1016/j.jtho.2018.05.013

98. Hirsch FR, McElhinny A, Stanforth D, et al. PD-L1 Immunohistochemistry assays for lung cancer: Results from phase 1 of the Blueprint PD-L1 IHC assay comparison project. *J Thorac Oncol*. 2017;12(2):208-222. doi:10.1016/j.jtho.2016.11.2228

99. Munari E, Rossi G, Zamboni G, et al. PD-L1 assays 22C3 and SP263 are not interchangeable in non-small cell lung cancer when considering clinically relevant cutoffs: An interclone evaluation by differently trained pathologists. *Am J Surg Path*. 2018;42(10):1384-1389. doi:10.1097/PAS.0000000000001105

100. Villaruz LC, Ancevski Hunter K, Kurland BF, Abberbock S, Herbst C, Dacic S. Comparison of PD-L1 immunohistochemistry assays and response to PD-1/L1 inhibitors in advanced non-small-cell lung cancer in clinical practice. *Histopathology*. 2019;74(2):269-275. doi:10.1111/his.13729

101. Adam J, Le Stang N, Rouquette I, et al. Multicenter harmonization study for PD-L1 IHC testing in non-small-cell lung cancer. *Ann Oncol*. 2018;29(4):953-958. doi:10.1093/annonc/mdy014

102. Kim H, Kwon HJ, Park SY, Park E, Chung J-H. PD-L1 immunohistochemical assays for assessment of therapeutic strategies involving immune checkpoint inhibitors in non-small cell lung cancer: A comparative study. *Oncotarget*. 2017;8(58):98524-98532. doi:10.18632/oncotarget.21567

103. Sughayer MA, Alnaimy F, Alsughayer AM, Qamhia N. Comparison of 22C3 pharmDx and SP263 assays to test PD-L1 expression in NSCLC. *Appl Immunohistochem Mol Morph*. 2019;27(9):663-666. doi:10.1097/PAI.0000000000000671

104. Grote HJ, Feng Z, Schlichting M, et al. Programmed death-ligand 1 immunohistochemistry assay comparison studies in NSCLC: Characterization of the 73-10 assay. *J Thorac Oncol*. 2020;15(8):1306-1316. doi:10.1016/j.jtho.2020.04.013

105. Singal G, Miller PG, Agarwala V, et al. Association of patient characteristics and tumor genomics with clinical outcomes among patients with non-small cell lung cancer using a clinicogenomic database. *JAMA*. 2019;321(14):1391-1399. doi:10.1001/jama.2019.3241

106. Aggarwal C, Thompson JC, Chien AL, et al. Baseline plasma tumor mutation burden predicts response to pembrolizumab-based therapy in patients with metastatic non-small cell lung cancer. *Clin Cancer Res*. 2020;26(10):2354-2361. doi:10.1158/1078-0432.CCR-19-3663

107. Liu L, Bai X, Wang J, et al. Combination of TMB and CNA stratifies prognostic and predictive responses to immunotherapy across metastatic cancer. *Clin Cancer Res*. 2019;25(24):7413-7423. doi:10.1158/1078-0432.CCR-19-0558

108. Chang H, Sasson A, Srinivasan S, et al. Bioinformatic methods and bridging of assay results for reliable tumor mutational burden assessment in non-small-cell lung cancer. *Mol Diagn Ther*. 2019;23(4):507-520. doi:10.1007/s40291-019-00408-y

109. Suzuki S, Haratani K, Hayashi H, et al. Association of tumour burden with the efficacy of programmed cell death-1/programmed cell death ligand-1 inhibitors for treatment-naive advanced non-small-cell lung cancer. *Eur J Cancer*. 2022;161:44-54. doi:10.1016/j.ejca.2021.11.011

110. Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol*. 2011;64(4):383-394. doi:10.1016/j.jclinepi.2010.04.026