

Fit-For-Purpose PD-L1 Biomarker Testing For Patient Selection in Immuno-Oncology: Guidelines For Clinical Laboratories From the Canadian Association of Pathologists-Association Canadienne Des Pathologistes (CAP-ACP)

Carol C. Cheung, MD, PhD, JD,† Penny Barnes, MD,‡ Gilbert Bigras, MD, PhD,§ Scott Boerner, MD,*† Jagdish Butany, MBBS, MS,*† Fiorella Calabrese, MD,¶ Christian Couture, MD, PhD,# Jean Deschenes, MD,** Hala El-Zimaity, MD, PhD,†† Gabor Fischer, MD, PhD,‡‡ Pierre O. Fiset, MD, PhD,§§ John Garratt, RT,||| Laurette Geldenhuys, MBBCH, FFPATH, MMED, FIAC, MAEd,‡ C. Blake Gilks, MD,||¶¶ Marius Ilie, MD, PhD,###*** Diana Ionescu, MD,††† Hyun J. Lim, PhD,‡‡‡ Lisa Manning, BSc,‡‡ Adnan Mansoor, MD,§§§ Robert Riddell, MD,†||| Catherine Ross, MD,¶¶¶ Sinchita Roy-Chowdhuri, MD, PhD,### Alan Spatz, MD, MSc,****†††‡‡‡‡‡ Paul E. Swanson, MD,§§§§||| Victor A. Tron, MD,†¶¶¶¶ Ming-Sound Tsao, MD,*† Hangjun Wang, MD,****†††‡‡‡‡‡ Zhaolin Xu, MD,‡ and Emina E. Torlakovic, MD, PhD,|||#####******

On behalf of the Canadian Association of Pathologists-Association Canadienne Des Pathologistes' National Standards Committee for High Complexity Testing (CAP-ACP NSCHCT)

Received for publication June 26, 2019; accepted June 28, 2019.

From the *Laboratory Medicine Program, Division of Pathology, University Health Network; †Department of Laboratory Medicine and Pathobiology, Faculty of Medicine, University of Toronto; ||||Department of Pathology and Laboratory Medicine, Mount Sinai Hospital; ¶¶¶¶Department of Laboratory Medicine, St. Michael's Hospital, Toronto; ††Dynacare Laboratories; ¶¶¶Juravinski Cancer Centre, Hamilton, ON; ‡Department of Pathology, Dalhousie University, Halifax, NS; §Department of Pathology, Cross Cancer Institute; **Department of Laboratory Medicine and Pathology, University of Alberta, Edmonton; §§§Department of Pathology and Laboratory Medicine, University of Calgary; |||||Calgary Laboratory Services, Calgary, AB; #Quebec Heart and Lung Institute, Université Laval, Quebec City; §§§Department of Pathology, McGill University; ‡‡‡‡Department of Pathology, McGill University; ****Division of Pathology and Molecular Genetics, McGill University Health Center; ††††Lady Davis Institute, Jewish General Hospital, Montreal, QC; ‡‡Department of Pathology, University of Manitoba, Winnipeg, MB; |||Canadian Immunohistochemistry Quality Control; ¶¶Department of Pathology and Laboratory Medicine, University of British Columbia; †††British Columbia Cancer Agency, Vancouver, BC; ‡‡‡Department of Community Health and Epidemiology; ****Department of Pathology and Laboratory Medicine, College of Medicine, University of Saskatchewan; #####Department of Pathology and Laboratory Medicine, Royal University Hospital, Saskatchewan Health Authority, Saskatoon, SK, Canada; ||Department of Cardiac, Thoracic, Vascular Sciences, and Public Health; ¶University of Padova Medical School, Padova, Italy; ##Laboratory of Clinical and Experimental Pathology; ***Hospital-Related Biobank (BB-0033-00025), Université Côte d'Azur, University Hospital Federation OncoAge, Hôpital Pasteur, Nice, France; ###Department of Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX; and §§§§Department of Pathology, University of Washington, School of Medicine, Seattle, WA. Supported by the Canadian Association of Pathologists—Canadienne Des Pathologistes (CAP-ACP), via unrestricted educational grants from AstraZeneca Canada, Bristol-Myers Squibb Canada, Merck Canada, and Roche Diagnostics. None of the industry sources of grant support had any role in the design of this document, selection of included sources, content analysis, discussion, conclusions, or in the decision of whether the paper would be submitted for publication, and where the paper would be submitted for publication. Precision Rx-Dx Inc. provided supplementary support to the National Standards Committee for program planning and organization. Jennifer Won and Heather Dow provided administrative support to the National Standards Committee.

All author declarations of potential conflicts of interest are in Appendix A.

Reprints: Emina E. Torlakovic, MD, PhD, Department of Pathology and Laboratory Medicine, Royal University Hospital, 103 Hospital Drive, Saskatoon, SK, Canada S7N 0W8 (e-mail: emina.torlakovic@saskhealthauthority.ca).

Supplemental Digital Content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website, www.appliedimmunohist.com.

Copyright © 2019 The Author(s). Published by Wolters Kluwer Health, Inc. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Abstract: Since 2014, programmed cell death protein 1 (PD-1)/programmed cell death ligand 1 (PD-L1) checkpoint inhibitors have been approved by various regulatory agencies for the treatment of multiple cancers including melanoma, lung cancer, urothelial carcinoma, renal cell carcinoma, head and neck cancer, classical Hodgkin lymphoma, colorectal cancer, gastroesophageal cancer, hepatocellular cancer, and other solid tumors. Of these approved drug/disease combinations, a subset also has regulatory agency-approved, commercially available companion/complementary diagnostic assays that were clinically validated using data from their corresponding clinical trials. The objective of this document is to provide evidence-based guidance to assist clinical laboratories in establishing fit-for-purpose PD-L1 biomarker assays that can accurately identify patients with specific tumor types who may respond to specific approved immunoncology therapies targeting the PD-1/PD-L1 checkpoint. These recommendations are issued as 38 Guideline Statements that address (i) assay development for surgical pathology and cytopathology specimens, (ii) reporting elements, and (iii) quality assurance (including validation/verification, internal quality assurance, and external quality assurance). The intent of this work is to provide recommendations that are relevant to any tumor type, are universally applicable and can be implemented by any clinical immunohistochemistry laboratory performing predictive PD-L1 immunohistochemistry testing.

Key Words: PD-L1, guidelines, biomarker, immunotherapy, quality assurance

(*Appl Immunohistochem Mol Morphol* 2019;27:699–714)

INTRODUCTION AND BACKGROUND

Programmed Cell Death Ligand 1 (PD-L1) Biology, Distribution in Normal Tissues, and Role in Immune Surveillance

The PD-L1 protein is encoded by the human *CD274* gene, located on the short arm of chromosome 9. It was first identified in 1999 based on a homology search of the putative functionally related molecules B7-1/2.¹ In 2000, programmed cell death protein 1 (PD-1) was identified as the receptor for the newly cloned PD-L1.² In vivo loss of function studies performed on knockout mice showed that PD-L1 negatively regulates T-cell function.³ Early in situ protein expression by immunohistochemistry (IHC) on frozen tissues showed expression in macrophages.⁴ Subsequently, IHC on formalin-fixed paraffin-embedded tissues showed that while PD-L1 expression in benign tissues may be limited to hematolymphoid cells and placenta, it is expressed by a number of different epithelial, mesenchymal, neuroectodermal, hematopoietic, lymphoid, and germ cell tumors.^{5,6}

The Evolution of PD-L1 as a Predictive Biomarker

PD-L1 has been evaluated as either a prognostic or a predictive biomarker in various malignant tumors.^{7–16} In contrast to previous immunotherapy checkpoint inhibitors, such as the anti-CTLA-4 drug ipilimumab,¹⁷ the use of the

anti-PD-1 inhibitor pembrolizumab in advanced non-small cell lung cancer (NSCLC) became the first immunotherapy drug to require PD-L1 biomarker testing (using a clinically validated United States' Food and Drug Administration (FDA)-approved IHC companion diagnostic assay) to determine patient eligibility based on results of the Keynote-001 trial, which demonstrated that the response to pembrolizumab is positively related to the level of PD-L1 expression.¹⁸ The Keynote 010 trial shifted the tumor proportion score (TPS) cutoff point for pembrolizumab as a second line treatment in NSCLC from 50% to 1%,¹⁹ whereas the Keynote 024 trial reintroduced the TPS \geq 50% in the first-line setting.²⁰ Subsequent to NSCLC, other tumors that have PD-L1 companion diagnostic testing requirements include gastric/gastroesophageal junction adenocarcinoma (pembrolizumab), cervical carcinoma (pembrolizumab), urothelial carcinoma (pembrolizumab and atezolizumab), breast cancer (atezolizumab), and head and neck squamous cell carcinoma (pembrolizumab). Tumors that have complementary PD-L1 testing assays approved by the FDA include NSCLC (nivolumab, atezolizumab), head and neck squamous cell carcinoma (nivolumab), urothelial carcinoma (nivolumab, durvalumab, atezolizumab), melanoma (nivolumab). Real-world outcomes of patients with metastatic NSCLC treated with PD-1 inhibitors in the first year following United States regulatory approval indicate that PD-L1 expression is associated with better overall survival.²¹

The PD-L1 Companion/Complementary Diagnostic Testing Landscape

Currently, there are 5 PD-1/PD-L1 inhibitors that are approved by the FDA for the treatment of various cancers including melanoma, lung cancer, urothelial carcinoma, renal cell carcinoma, head and neck cancer, classical Hodgkin lymphoma, colorectal cancer, gastroesophageal cancer, hepatocellular cancer, and other solid tumors.²² Of these approved drug/disease combinations, a subset has FDA-approved, commercially available companion or complementary diagnostic assays that were clinically validated using data from their corresponding clinical trials or from related clinical studies. Although PD-L1 biomarker status to determine eligibility for pembrolizumab in NSCLC has significance for clinical laboratories as being the first companion diagnostic assay in immuno-oncology,¹⁸ the overall landscape suggests that PD-L1 testing will continue to expand.^{23–28} This likelihood is augmented by the possibility that (irrespective of the FDA) regulatory or funding agencies in countries other than the United States, may require demonstration of “PD-L1 positivity” by a properly validated IHC assay to approve or fund the use of anti-PD-L1/PD-1 therapies for specific patients as higher PD-L1 expression has been positively associated with outcome even in those drug-disease indications that currently do not require PD-L1 testing.^{29–34}

Implementing PD-L1 Testing: Challenges For Clinical Laboratories

There are multiple different commercially available PD-L1 IHC companion/complementary diagnostic assays (PD-L1 IHC Kits). These PD-L1 assays were developed

and clinically validated by clinical trials^{35–41} for different drug-disease combinations, using different PD-L1 primary antibody (Ab) clones on different IHC platforms with different IHC protocols and requiring different readout criteria for what is considered to be “positive” in different disease contexts. PD-L1 IHC Kits are reasonably considered reference standard assays for their respective drug-disease indications based on the Clinical and Laboratory Standards Institute (CLSI) guidelines for determining the diagnostic accuracy of qualitative assays. However, because the analytical sensitivity and specificity (including the acceptable threshold range relevant to low limit of detection) are not precisely defined for these assays, the question of whether or not the different assays can be used in place of one another (ie, are interchangeable) or whether laboratory-developed tests (LDTs) can be successfully validated, inevitably arises.

These questions represent significant challenges to clinical laboratories charged with the task of providing meaningful, reliable, and informative PD-L1 testing results in an environment where such results have an impact on eligibility for therapy and where oncologists and patients, rightfully, have high expectations of test accuracy. From the laboratory perspective, the challenges in establishing fit-for-purpose PD-L1 testing are multifaceted with biological, technical, interpretive, and regulatory factors coming together to form what amounts to an empirical gauntlet for laboratories. As such, there is an emerging need for guidance in this matter, a clear roadmap for the selection, validation/verification and reporting of PD-L1 IHC assays that are fit-for-purpose and evidence-based.

OBJECTIVE

The scope of these Canadian Association of Pathologists-Association Canadienne Des Pathologistes (CAP-ACP) recommendations is to provide guidance to assist clinical laboratories in the setup and implementation of fit-for-purpose PD-L1 biomarker assays that can accurately identify patients with specific tumor types who may respond to specific Health Canada-approved immuno-oncology therapies targeting the PD-1/PD-L1 checkpoint. These recommendations are issued as guideline statements (GSs) that address (i) assay selection and development for surgical pathology and cytopathology specimens, (ii) reporting elements, and (iii) quality assurance [including validation/verification, internal quality assurance, and external quality assurance (EQA)]. The intent of this work is to provide recommendations that are relevant to any tumor type, are universally applicable and can be implemented by any clinical IHC laboratory performing predictive PD-L1 IHC testing.

These recommendations do not apply to (and should not inform) drug selection by oncologists for any patient populations; rather, drug selection for specific disease indications is relevant to these recommendations only to the extent that the drug-disease combinations impact the selection and implementation of an appropriate fit-for-purpose PD-L1 testing strategy. Importantly, these recommendations

are not designed to address specific IHC protocols or protocol components. Similarly, they do not address the issues of whether PD-L1 testing should be performed or the context(s) in which PD-L1 testing may be informative.

The GSs address 5 key questions:

- (1) Which PD-L1 assay(s) should be used to predict response to anti-PD-1/PD-L1 immunotherapies?
- (2) What is the quality of statistical methodologies employed to evaluate PD-L1 assay performance in interchangeability assessments?
- (3) Were specific diagnostic assays (IHC protocol conditions and specific readout) used and stated by clinical trials where a specific drug and a specific disease were evaluated?
- (4) How should the results of predictive PD-L1 assays be reported?
- (5) What measures/practices are necessary to ensure the quality of PD-L1 testing for patient selection in immunotherapy?

METHODOLOGY

The methodology is detailed in the Supplementary Files (Supplemental Digital Content 1, <http://links.lww.com/AIMM/A241>). Systematic and targeted review of published evidence, grading of evidence, development of recommendations, and grading of recommendations was done according to published guidelines with some modifications^{42–53} in relation to “implementability” of the recommendations to daily practice in clinical IHC laboratories.⁵⁴ A systematic review of published literature for key questions 1 and 2 resulted in the initiation of a meta-analysis of published test comparisons with emphasis on diagnostic accuracy being a principal criterion for test interchangeability; see Supplementary Materials—Key Questions (Supplemental Digital Content 1, <http://links.lww.com/AIMM/A241>).

HOW TO USE THIS DOCUMENT

The main outputs of this initiative are the 38 GSs, which are found in the section below and summarized in Table 1. Each GS is accompanied by an explanatory note where the authors felt it was necessary.

TERMINOLOGY

Accuracy

The closeness of a single result of a measurement and a true value.^{55,56}

Biomarker

A physiological analyte that is objectively measured and evaluated as an indicator of normal biological and pathogenic processes or pharmacological responses to a specified therapeutic intervention.⁵⁶

Candidate Assay

An assay under evaluation also referred to as an “index assay.”^{55,57}

TABLE 1. Guideline Statements

Number	Guideline Statement	Strength of Recommendation
<i>1.1. Recommendations for Assay and Sample Selection - Assay Selection</i>		
GS01	Selection of a PD-L1 IHC assay as a predictive biomarker for anti-PD-1/PD-L1 treatment must be made in a fit-for-purpose manner in accordance with the 3D approach (Drug-Disease-Diagnostic assay).	Strong recommendation
GS02	Health Canada-approved, FDA-approved, and CE-marked PD-L1 IHC Kits that are validated in clinical trials for specific purposes must be used for those specific purposes.	Strong recommendation
GS03	Health Canada-approved, FDA-approved, and CE-marked PD-L1 IHC Kits that are validated in clinical trials for specific purposes, when not used for these specific purposes, but used for a different purpose (ie, in accordance with the 3D approach for a different disease and/or different drug), must be considered as LDTs and are not to be considered as Health Canada-approved, FDA-approved, or CE-marked for this new purpose.	Strong recommendation
GS04	PD-L1 IHC LDTs that are properly validated for a specific purpose according to Section 3.1 (Recommendations for Quality Assurance - Validation) may be used for that purpose, even if not Health Canada-approved, FDA-approved, or CE-marked for this new purpose. This applies whether they are developed with the same or a different primary Ab (clone or polyclonal Ab) compared with IVDs marketed for the same purpose.	Strong recommendation
GS05	Selection of a "replacement assay" for any given purpose (as defined by the 3D approach) should be based on demonstrated diagnostic test accuracy (sensitivity and specificity) against a designated reference standard (or other diagnostic accuracy criteria) for that specific purpose; neither clone nor readout need to be identical, but high sensitivity and specificity (see Section 3.1; Recommendations for Quality Assurance - Validation) should be achieved for direct or indirect clinical validation by the selected replacement (candidate) assay.	Recommendation
GS06	Published results of properly conducted studies of direct or indirect clinical validation and/or meta-analyses of such studies, that address diagnostic accuracy of candidate assays, must guide selection of potential "replacement assays" for any specific purpose (eg, both Dako PD-L1 IHC 28-8 pharmDx and VENTANA PD-L1 (SP263) showed acceptable diagnostic accuracy for the 50% TPS cutoff for PD-L1 IHC 22C3, but not for the 1% cutoff).	Strong recommendation
GS07	For Health Canada-approved, FDA-approved, and CE-marked PD-L1 IHC Kits, the clinical laboratory should validate the assay as an LDT when any preanalytical parameters fall outside of what is recommended in the Kit's instructions for use	Recommendation
GS08	For LDTs that employ the same primary antibody clone as a Health Canada-approved, FDA-approved, and CE-marked PD-L1 IHC Kit, the same pre-analytical caveats found in the corresponding Health Canada-approved, FDA-approved, and CE-marked PD-L1 IHC Kit apply.	Expert opinion
<i>1.2. Recommendations for Assay and Sample Selection - Sample Selection.</i>		
GS09	If more than one tissue block is available for a given tumor, the most representative sample should be tested. More than one block may be tested when the reporting pathologist determines that such additional testing is necessary to establish the PD-L1 status of the tumor. If additional blocks from the same sample are tested, the results from all tested blocks should be combined as if it were present in a single paraffin block.	Recommendation
GS10	For samples where the initial level contains less than the minimum required number of tumor (or other target) cells as defined by the interpretive manual for a specific assay and purpose, testing of additional levels may be performed with extra tissue containing additional tumor cells to reach the required threshold. If additional levels are tested, all tissue from all levels may be combined as if it were present in a single level.	Expert opinion
GS11	If > 1 sample is available at the same time (eg, biopsies from two different sites, or biopsy and cytology samples, or primary and corresponding metastatic tumors), the most representative sample should be tested. More than one sample may be tested when the reporting pathologist determines that such additional testing is necessary to establish the PD-L1 status of the tumor.	Recommendation
GS12	It is acceptable to test a tumor sample at the time of consideration for anti-PD-1/PD-L1 immunotherapy even if an older tumor sample was previously tested.	Recommendation
GS13	The use of Health Canada-approved, FDA-approved, or CE-marked PD-L1 IHC Kits that were validated for formalin-fixed paraffin-embedded biopsy samples, and not specifically validated for cytology samples, may be used for cytology samples if (i) they were processed according to the same pre-analytical conditions as required by the PD-L1 IHC Kits, and (ii) the readout is compatible with the type of cytology samples considering the lack of tissue architecture.	Recommendation
GS14	For the use of PD-L1 LDTs on cytology samples, GS30 applies.	Expert opinion
GS15	For samples where identification of appropriate cell types is problematic (eg, tumor cell vs. macrophage in NSCLC), multiplexing (eg, double IHC) may be used, not instead of, but in addition to the usual PD-L1 assay that is performed by the laboratory, so long as the multiplexed assay has undergone appropriate technical validation.	Expert opinion
<i>2. Recommendations for Assay/Test Reporting - Reporting of Predictive PD-L1 Assays</i>		
GS16	Pathologists issuing reports for PD-L1 testing should successfully undergo training that is (i) endorsed by the manufacturer of an approved PD-L1 IHC Kit, or (ii) endorsed by a professional organization (eg, Canadian Association of Pathologists, International Quality Network for Pathology, etc.), or (iii) provided by a pathologist who has undergone training as indicated in (i) or (ii) and in a train-the-trainer format. The optimal goal for such training should be to achieve 90% sensitivity and 90% specificity of the previously validated readout.	Expert opinion
GS17	Reporting of computer-assisted PD-L1 readout (image analysis) for the purpose of patient selection for immunotherapy is not recommended at this time.	Expert opinion
GS18	A structured (synoptic) format should be used for PD-L1 reporting. Searchable formats are recommended to facilitate audit of institutional or pathologist-specific prevalence.	Recommendation
GS19	In the comments section of the report, the following information should be included: (i) the drug and disease for which the assay was run, and (ii) the assay that was performed, and on what platform, (iii) the clinically relevant cutoff point for the	Recommendation

TABLE 1. (continued)

Number	Guideline Statement	Strength of Recommendation
	3D combination, (iv) the readout parameter being reported (eg, Tumour Proportion Score, Combined Positive Score, Immune Cells), and (v) the clinically relevant cutoff category within which the result falls.	
GS20	If the specimen was inadequate for testing, this may be stated in the report along with the reason why the specimen was inadequate.	Expert opinion
GS21	Results with cytology samples may be reported in the same manner recommended for solid tissue samples (see GS19)	Expert opinion
GS22	Results with fresh-cut sections from archived paraffin blocks or previously-cut and stored unstained sections that are older than the PD-L1 IHC Kit manufacturer's recommendation may be reported only if positive and it is the pathologist's assessment that, based on presented evidence, the result is not false-positive (eg, pattern of staining being membranous, appropriate internal and external controls are present and show expected range of positive and negative results). "Positive result" is here defined as a positive result for a specific purpose and a specific patient context (eg, the staining of inflammatory cells has no relevance for some PD-L1 assays while for others, they are the principal cell type assessed to determine whether the sample is to be designated as being a "positive" result). Therefore, some positivity identified in a lesion does not qualify the results as reportable. Negative results may be reported as "no result" and potential for false-negative results with older samples may be included in the comment section.	Expert opinion
	<i>3.1. Recommendations for Quality Assurance - Validation</i>	
GS23	Validation of PD-L1 IHC predictive assays must be fit-for-purpose in accordance with the 3D (Drug-Disease-Diagnostic assay) approach.	Strong recommendation
GS24	When a laboratory decides that an LDT will be used instead of a regulatory agency-approved IHC kit for the same purpose, the IHC laboratory must provide evidence of successful methodology transfer. This must include (i) technical validation and (ii) indirect clinical validation by using the relevant reference standard where not already published or established (ie, the regulatory agency-approved kit for the same purpose).	Strong recommendation
GS25	Indirect clinical validation can be performed when there is a recognized reference standard or gold standard that has been validated in a published clinical trial for that specific purpose. Samples for indirect clinical validation should include at least 50 randomly selected, clinically relevant positive cases, and at least 50 negative cases identified in the screening process.	Recommendation
GS26	PD-L1 IHC LDTs that have been clinically validated or indirectly clinically validated in published literature must still be technically validated by the clinical IHC laboratory before being put in use as a PD-L1 predictive biomarker assay.	Strong recommendation
GS27	Health Canada-approved, FDA-approved, and CE-marked PD-L1 IHC Kits are already directly or indirectly clinically validated in clinical trials; however, their analytical performance must be verified according to manufacturer's instructions before use. If no specific instructions are provided by the manufacturer, general recommendations for Tier 1 technical validation apply (analytical sensitivity, analytical specificity, and reproducibility) by using iCAPCs or iCAPCs-like tissues/calibrators with at least 3 successful, independent verification runs.	Strong recommendation
GS28	As per recommendations above, if multiplex assays (eg, double IHC staining for PD-L1 and TTF-1) are to be used, they must have undergone technical validation. The possible impact of multiplexing on diagnostic accuracy of each biomarker must be assessed against single IHC staining using a Health Canada-approved, FDA-approved, or CE-marked IHC kit or other fit-for-purpose designated reference standard.	Strong recommendation
GS29	The PD-L1 IHC readout must be validated for diagnostic accuracy (readout accuracy) by each pathologist reporting PD-L1 IHC results as a categorical variable; in addition, validation of readout precision must also be conducted when applicable (eg, if reporting exact % positive cells as a continuous variable). This applies for both regulatory agency-approved IHC kits and LDTs.	Strong recommendation
GS30	Laboratories using regulatory agency-approved PD-L1 IHC Kits in a manner that deviates from the manufacturer's instructions must validate the deviation.	Strong recommendation
	<i>3.2. Recommendations for Quality Assurance - Quality Control</i>	
GS31	Laboratories must run external on-slide controls for every slide.	Strong recommendation
GS32	The minimum composition of external controls for FDA (or other regulatory body) approved PD-L1 IHC Kits must be in compliance with the Kit manufacturer's instructions.	Strong recommendation
GS33	The minimum composition of external on-slide controls for LDTs ideally will mirror the requirements of the corresponding fit-for-purpose FDA (or other regulatory body)-approved kit.	Expert opinion
GS34	Where not already included as part of the required minimum composition, external on-slide controls for PD-L1 testing may include the following elements: (i) lesional tissue showing positivity for PD-L1 in a specified cell population, (ii) lesional tissue showing negativity for PD-L1 in a specified cell population, (iii) low expressor tissue in accordance with iCAPCs-like principles (eg, macrophages in germinal centers of tonsil).	Expert opinion
GS35	The laboratory should assess the external on-slide control by light microscopy before a slide being released from the laboratory. When external on-slide controls demonstrate either unacceptable background or lack of signal in any positive control tissue, such IHC slides should be failed by the laboratory.	Recommendation
	<i>3.3. Recommendations for Quality Assurance - External Quality Assurance - Proficiency Testing</i>	
GS36	All laboratories performing PD-L1 testing for patient selection should participate in fit-for-purpose PT at least every six months where available and required by relevant regulatory bodies.	Recommendation
GS37	Laboratories should select a PT program that ideally informs participating laboratories on the accuracy of the PD-L1 IHC protocol and the pathologists' readout as determined by central expert assessment. If no such program is available/accessible, inter-laboratory exchange of samples and information with a provincial or national reference laboratory should be performed.	Recommendation
GS38	Participation in PT, even when PT is fit-for-purpose, is not acceptable as a substitute for clinical and technical validation.	Strong recommendation

Characteristics of Validation

Validation of laboratory assays cannot be performed without defining validation characteristics. There are 4 essential characteristics of validation of laboratory assays including (i) sphere of validation, (ii) type of validation, (iii) scope of validation, and (iv) tier of validation. Tiers apply mostly to technical validation.⁵⁸

Clinical Validation of Predictive Biomarker

The *process* of demonstrating how robustly and reliably the biomarker result predict the clinical outcome of interest. This is also termed as “qualification of predictive biomarker” (see below).⁵⁹

Clinical Validity (Synonym—Clinical Utility)

The assay’s *ability* to add value to patient management decision-making compared with current practices.⁶⁰

Companion Diagnostic IHC Assay For PD-L1

An IHC assay that provides information for the effective use of a corresponding anti-PD-1/PD-L1 therapeutic product, linked to the specific drug within its approved labeling.⁶¹

Comparator Assay

An assay that was designated as the true value also referred to as “reference standard” or “diagnostic accuracy criteria.”^{55,57}

Complementary Diagnostic IHC Assay For PD-L1

An IHC assay that can aid in therapeutic decision-making for anti-PD-1/PD-L1 therapeutic products but is not required before prescribing a drug.⁶²

Diagnostic Accuracy

The extent of agreement between a candidate assay (ie, index assay) and a comparator assay (ie, reference standard or other diagnostic accuracy criteria).^{55,57}

Diagnostic Accuracy Criteria

The best currently available criteria for establishing the presence or absence of the condition, event, or characteristic.⁵⁵

Diagnostic Biomarker

A biomarker that is used to identify disease.⁶³

Diagnostic Sensitivity

The proportion of those with the target condition (as defined by a reference standard) who test positive with a candidate test.^{55,57,64}

Diagnostic Specificity

The proportion of those without the target condition (as defined by a reference standard) who test negative with a candidate assay.^{55,57,64}

Diagnostic Validation of a Biomarker

The *process* of demonstrating how robustly and reliably the assay results correlate with the diagnosis of interest.⁵⁸

EQA Accuracy

For PD-L1, this is demonstrated by using 20 positive and 20 negative cases⁶⁵ (based on a fit-for-purpose cutoff point) of tumors for which the PD-L1 biomarker will be used and the results compared with those obtained by another laboratory that is successfully using an already validated assay; positive cases are selected as such to span the entire reportable range of PD-L1 expression and positive cases usually originate from the institutional tissue archive. The aim of EQA accuracy assessment is to demonstrate that the LDT protocol performs as it should in a specific type of tumor (with specificity for predefined types of cells/tissues, appropriate subcellular signal localization, etc.).⁶⁶ This is not to be confused with indirect clinical validation in which a larger number of randomly selected cases is used and evaluated against a reference standard IHC assay (see below).^{55,57}

Indirect Clinical Validation of a Predictive Biomarker Assay

Validation of diagnostic accuracy of a biomarker assay against the results of a designated, previously clinically validated reference standard assay or “diagnostic accuracy criteria” (see above) or designated reference standard biomarker assay. The previously validated biomarker assay may or may not be regulatory agency-approved and it may or may not be employing the same methodology (eg, IHC, but also fluorescence in situ hybridization, etc.), but it must be already qualified/validated in a prospective clinical trial. Indirect clinical validation demonstrates that the assay in question (candidate assay) has identical or nearly identical diagnostic sensitivity and specificity as the reference standard assay or “diagnostic accuracy criteria” (see above) or designated reference standard assay (comparator assay), where the comparator assay has already established link(s) to clinical outcomes; it attempts to answer the question of whether (or to what degree) a candidate assay may identify the same patients (as being “positive” or “negative”) as the comparator assay.

“Interchangeable” PD-L1 IHC Predictive Biomarker Assay

This definition is adapted to be analogous to the use of the term “interchangeable” in drug development and approval. It refers to PD-L1 IHC assays that have demonstrated essentially identical performance in clinical trials for the same disease and the same drug. Demonstration of “correlation,” “similarity,” or “overall agreement” must not be interpreted as a demonstration of identical clinical outcomes, but rather technical performance. Indirect clinical validation is not sufficient for the clinical qualification of a biomarker assay and cannot be used as a basis for designating a biomarker assay as “interchangeable.”⁵⁹ This applies to both, FDA-approved kits as well as LDTs.

LDT

An LDT, as it pertains to this document, is any IHC assay designed to detect and report the expression of PD-L1 protein in tumor cells and/or immune cells for

predicting potential response to a regulatory agency–approved PD-L1/PD-1 checkpoint inhibitor with at least one of the following characteristics: (1) A testing laboratory develops and validates a PD-L1 IHC assay from first principles using separately purchased, commercially available components (aka. “de novo LDT”); (2) A testing laboratory adds/subtracts/modifies any manufacturer-specified preanalytical, analytical, or postanalytical component/aspect of a commercially available, regulatory agency–approved PD-L1 IHC assay/in vitro diagnostic device (aka. “kit-derived LDT”); (3) A testing laboratory performs PD-L1 testing using a commercially available, regulatory agency–approved PD-L1 IHC assay/in vitro diagnostic device in accordance with the manufacturer’s specifications but for a purpose other than that intended by the manufacturer (aka. “kit-derived LDT”). This definition of LDT is adapted from the Canadian Standards Association/Standards Council of Canada’s standard Z316.8-18: Requirements for the design, development, and validation of LDTs.⁶⁷

PD-L1 IHC Assay

Any IHC assay where the purpose is to demonstrate expression of the PD-L1 protein.

PD-L1 IHC Biomarker

Any PD-L1 IHC assay where both IHC protocol and IHC readout are fit-for-purpose based on demonstrated evidence.

Predictive Biomarker

A biomarker used to identify individuals who are more likely than similar individuals without the biomarker to experience a favorable or unfavorable effect from exposure to a medical product or an environmental agent.⁶⁸

Prognostic Biomarker

A biomarker that provides information on the likely patient health outcome (eg, disease recurrence or progression) regardless of the treatment.⁶⁸

Qualification

A conclusion, based on a formal regulatory process, that within the stated context of use, a medical product development tool can be relied upon to have a specific interpretation and application in medical product development and regulatory review⁶³; when applied to a predictive biomarker, it refers to “clinical validation.”^{68–70}

Regulatory Agency–Approved IHC Kit/Assay

An IHC kit/assay that was approved by a regulatory agency for a specific purpose.

Replacement Assay

See “Candidate assay.”

“Repurposed” PD-L1 IHC Predictive Biomarker Assay

Analogous to repurposed drugs, repurposed predictive biomarker assays those that have been originally developed and qualified in a clinical trial for 1 purpose

(specific drug and disease) but were also qualified in a different clinical trial for a different purpose.^{71,72} Such a repurposed PD-L1 IHC biomarker assay is expected to have identical IHC protocol conditions, but may have different readout and different reporting requirements, for example, the PD-L1 IHC 22C3 pharmDx assay for NSCLC was repurposed as a biomarker assay for PD-L1 detection in gastric cancer for immunotherapy with pembrolizumab; a different readout was designated to this assay for this new purpose.⁷³

Revalidation (technical/analytical)

In clinical IHC, technical/analytical revalidation is divided into primary, secondary, and tertiary revalidation depending on the trigger that initiated revalidation.⁵⁸

Scope of Technical/Analytical Validation

In clinical IHC, the scope of technical/analytical validation is divided into initial validation and revalidation.⁵⁸

Spheres of Validation

IHC validation has the following spheres: clinical, indirect clinical/diagnostic, technical/analytical.⁵⁸

Technical/Analytical Validation of IHC Assay

Assessment of performance characteristics of an assay, including analytical sensitivity, analytical specificity, analytical reproducibility, EQA and/or non-IHC methodology accuracy, reportable range, extended analytical specificity, and preanalytical robustness.^{58,59,66,74}

Tiers of Technical/Analytical Validation

Tier 1 (synonym: “verification”; including analytical sensitivity, analytical specificity, and reproducibility), tier 2 (including EQA or non-IHC methodology accuracy, reportable range, and extended analytical specificity), tier 3 (validation for preanalytical robustness including analytical sensitivity, analytical specificity for different clinically and institutionally applicable preanalytical conditions, reagents, and times).⁵⁸ Verification is typically performed by using control materials such as immunohistochemistry critical assay performance controls (iCAPCs).⁶⁴

Types of Validation

IHC validation type refers to the subject of validation. This includes validation of reagents (eg, buffer, primary Ab validation), lot-to-lot validation of primary Ab, validation of IHC protocol, validation of pathologist’s readout, validation of instruments, etc.⁵⁸

Validation

The process of assay/test validation establishes the clinical and analytical performance characteristics of an assay/test as well as the assay/test limitations. This involves confirmation, through the provision of objective evidence, that the requirements for a specific intended use or application have been fulfilled.^{58,75}

Verification

Verification is performed to ensure that the laboratory can meet or exceed the Health Canada-approved,

FDA-approved, or CE-marked performance characteristics established by the assay/test manufacturer; this involves confirmation, through the provision of objective evidence, that specified requirements have been fulfilled.⁷⁵ Verification is also known as tier 1 technical/analytical validation (see above).⁵⁸

RESULTS

A total of 38 GSs were generated along 3 broad streams: (1) Recommendations for assay and sample selection, (2) Recommendations for assay/test reporting, and (3) Recommendations for quality assurance.

Recommendations for Assay and Sample Selection

Assay Selection

GS01. Strong recommendation—*Selection of a PD-L1 IHC assay as a predictive biomarker for anti-PD-1/PD-L1 treatment must be made in a fit-for-purpose manner in accordance with the 3D approach (Drug-Disease-Diagnostic assay).*^{18,23,29–34,76–86}

Explanatory note: In biomarker-driven companion/complementary diagnostic testing, “purpose” is tripartite and consists of (i) the Drug, (ii) the Disease, and (iii) the Diagnostic assay. What links the 3D’s of purpose together is the clinical trial where the efficacy of the Drug was established in participants with a specific Disease and where biomarker testing results with a specific Diagnostic assay on biospecimens from trial participants could successfully separate participants who showed clinical response to the Drug from participants who did not show clinical response to the Drug. A Diagnostic assay that has been validated with biospecimens from responders and nonresponders in a clinical trial designed to assess the efficacy of a drug in a specific disease, is considered to be “clinically validated” or “qualified.”

GS02. Strong recommendation—*Health Canada-approved, FDA-approved, and CE-marked PD-L1 IHC Kits that are validated in clinical trials for specific purposes must be used for those specific purposes.*^{76–78,86–90}

Explanatory note: “PD-L1 IHC Kit” is a commercially manufactured and marketed Diagnostic assay that uses IHC to detect certain expression patterns of PD-L1 protein and that was clinically validated (see GS01) with biospecimens from responders and nonresponders of a clinical trial for a specific drug in a specific disease population.

GS03. Strong recommendation—*Health Canada-approved, FDA-approved, and CE-marked PD-L1 IHC Kits that are validated in clinical trials for specific purposes, when not used for these specific purposes, but used for a different purpose (ie, in accordance with the 3D approach for a different disease and/or different drug), must be considered as LDTs and are not to be considered as Health Canada-approved, FDA-approved, or CE-marked for this new purpose.*^{86–88,90–93}

Explanatory note: If Diagnostic assay #1, which was clinically validated with results from its clinical trial (ie, clinical trial #1) that assessed the efficacy of Drug #1 in Disease #1, is used to predict potential response to Drug

#2 in Disease #2, then Diagnostic assay #1 is considered a LDT in the context of Drug #2 in Disease #2. In the context of Drug #1 in Disease #1 though, Diagnostic assay #1 is a fit-for-purpose companion/complementary diagnostic assay (eg, PD-L1 IHC Kit #1). Such “kit-derived LDTs” that were also indirectly clinically validated for some purposes (that were not included in the kit label) in published literature, may be acceptable for those limited additional applications. However, just like any other LDT, “kit-derived LDTs” need to be technically validated by the laboratory that will be performing the PD-L1 testing before being put in clinical use.

GS04. Strong recommendation—*PD-L1 IHC LDTs that are properly validated for a specific purpose according to Section 3.1 (Recommendations for Quality Assurance - Validation) may be used for that purpose, even if not Health Canada-approved, FDA-approved, or CE-marked for this new purpose. This applies, whether they are developed with the same or a different primary Ab (clone or polyclonal Ab) compared with IVDs marketed for the same purpose.*^{58,59,65,94}

Explanatory note: If a biomarker test is required or desired to predict for potential response to a specific approved Drug in a specific Disease AND the laboratory does not wish to or is not able to use the specific PD-L1 IHC Kit that was clinically validated by the associated clinical trial, then the laboratory may use an LDT so long as the LDT is validated in a fit-for-purpose manner in accordance with Section 3.1 (Recommendations for Quality Assurance - Validation). An LDT may be developed using any anti-PD-L1 primary Ab clone as long as the LDT is properly validated. Well-designed, fit-for-purpose (properly validated) de novo LDTs are favored over kit-derived LDTs (see GS03). See GS26 for the definition of a “properly validated LDT.”

GS05. Recommendation—*Selection of a “replacement assay” for any given purpose (as defined by the 3D approach) should be based on demonstrated diagnostic test accuracy (sensitivity and specificity) against a designated reference standard (or other diagnostic accuracy criteria) for that specific purpose; neither clone nor readout need to be identical, but high sensitivity and specificity (see Section 3.1; Recommendations for Quality Assurance - Validation) should be achieved for direct or indirect clinical validation by the selected replacement (candidate) assay.*⁵⁵

Explanatory note: This statement reiterates that a properly validated PD-L1 IHC LDT is one that shows evidence of success for both technical validation and clinical validation (direct or indirect). It further elaborates that the evidence generated by indirect clinical validation is proof of diagnostic accuracy, which consists of 2 elements: diagnostic sensitivity and diagnostic specificity. Diagnostic sensitivity and specificity should both be $\geq 90\%$. For indirect clinical validation, this means that at least 90% of cases that were read as being positive based on results generated by the reference standard PD-L1 IHC Kit must also be read as being positive based on results generated by the corresponding PD-L1 IHC LDT. Similarly, at least 90% of cases that were read as being negative based on results generated by the reference standard PD-L1 IHC

Kit must also be read as being negative based on results generated by the corresponding PD-L1 IHC LDT. The tissues for indirect clinical validation may originate from the institution that is performing the indirect clinical validation for a newly developed LDT, but it also may originate from a proficiency testing (PT) program's reference laboratory as well as from multiinstitutional sources.

GS06. Strong recommendation—Published results of properly conducted studies of direct or indirect clinical validation and/or meta-analyses of such studies, that address diagnostic accuracy of candidate assays, must guide selection of potential “replacement assays” for any specific purpose (eg, both Dako PD-L1 IHC 28-8 pharmDx and VENTANA PD-L1 (SP263) showed acceptable diagnostic accuracy for the 50% TPS cut-off for PD-L1 IHC 22C3, but not for the 1% cut-off).⁹⁰

Explanatory note: A systematic review and meta-analysis for PD-L1 replacement assays indicated that the evidence does not support interchangeability of the assays when based on diagnostic accuracy.⁹⁰

GS07. Recommendation—For Health Canada-approved, FDA-approved, and CE-marked PD-L1 IHC Kits, the clinical laboratory should validate the assay as an LDT when any pre-analytical parameters fall outside of what is recommended in the Kit's instructions for use.^{58,59,87,88,92,94,95}

Explanatory note: Validation refers to both clinical validation (direct or indirect) and technical validation. As explained above, any parameter outside of the declared label of a Kit renders the assay to an LDT and therefore, at a minimum it requires indirect clinical validation and technical validation. See also GS30.

GS08. Expert opinion—For LDTs that employ the same primary antibody clone as a Health Canada-approved, FDA-approved, and CE-marked PD-L1 IHC Kit, the same pre-analytical caveats found in the corresponding Health Canada-approved, FDA-approved, and CE-marked PD-L1 IHC Kit apply.^{59,96}

Explanatory note: Every monoclonal primary Ab is designed for a unique epitope that has its own biochemical characteristics including preanalytical robustness. When pre-analytical robustness of 1 primary Ab (eg, 22C3 clone) is tested, these results do not necessarily apply to other clones developed for different epitopes of the same molecule (eg, 28-8, SP142, and SP263, all of which are developed to detect the PD-L1 molecule but bind to different epitopes).

Sample Selection

GS09. Recommendation—If more than one tissue block is available for a given tumor, the most representative sample should be tested.^{97–122} More than one block may be tested when the reporting pathologist determines that such additional testing is necessary to establish the PD-L1 status of the tumor. If additional blocks from the same sample are tested, the results from all tested blocks should be combined as if it were present in a single paraffin block.

Explanatory note: Depending on the pathologist's assessment, testing of additional paraffin blocks may be required to establish the PD-L1 status of the tumor.

GS10. Expert opinion—For samples where the initial level contains less than the minimum required number of tumor (or other target) cells as defined by the interpretive manual for a specific assay and purpose, testing of additional levels may be performed with extra tissue containing additional tumor cells to reach the required threshold. If additional levels are tested, all tissue from all levels may be combined as if it were present in a single level.

No Explanatory note.

GS11. Recommendation—If more than one sample is available at the same time (eg, biopsies from two different sites, or biopsy and cytology samples, or primary and corresponding metastatic tumors), the most representative sample should be tested. More than one sample may be tested when the reporting pathologist determines that such additional testing is necessary to establish the PD-L1 status of the tumor.^{97–122}

No Explanatory note.

GS12. Recommendation—It is acceptable to test a tumor sample at the time of consideration for anti-PD-1/PD-L1 immunotherapy even if an older tumor sample was previously tested.¹²²

No Explanatory note.

GS13. Recommendation—The use of Health Canada-approved, FDA-approved, or CE-marked PD-L1 IHC Kits that were validated for formalin-fixed paraffin-embedded biopsy samples, and not specifically validated for cytology samples, may be used for cytology samples if (i) they were processed according to the same pre-analytical conditions as required by the PD-L1 IHC Kits, and (ii) the readout is compatible with the type of cytology samples considering the lack of tissue architecture.^{108,123–130}

Explanatory note: PD-L1 IHC Kits are typically validated for formalin-fixed paraffin-embedded surgical pathology samples and not for “cytology” samples. This would exclude smears since regardless of fixative, smears are typically not paraffin-embedded. However, cytology specimens that are immediately fixed in 10% neutral buffered formalin and then spun down into a cell pellet that is subsequently processed/paraffin-embedded (similar to typical surgical pathology specimens) and where tumor cells can be clearly distinguished from inflammatory cells and other cells (eg, mesothelial cells), are essentially small biopsies and may be tested. However, if the laboratory intends to perform PD-L1 testing on i) smears (regardless of fixative) or ii) cytologic specimens not fixed in 10% neutral buffered formalin, then the PD-L1 IHC Kit becomes an LDT in the context of the nonqualifying “cytology” specimen-type and indirect clinical validation is required.

GS14. Expert opinion—For the use of PD-L1 LDTs on cytology samples,^{58,65,76,90} GS30 applies.

Explanatory note: PD-L1 IHC LDTs are considered properly validated for cytology samples when evidence of clinical validation (direct or indirect) and technical validation is provided and is relevant to the cytology samples on which the laboratory plans to perform testing. Also, see GS26.

GS15. Expert opinion—For samples where identification of appropriate cell types is problematic (eg, tumor

cell vs. macrophage in NSCLC), multiplexing (eg, double IHC) may be used, not instead of, but in addition to the usual PD-L1 assay that is performed by the laboratory, so long as the multiplexed assay has undergone appropriate technical validation.^{58,65}

No Explanatory note.

Recommendations for Assay/Test Reporting

Reporting of Predictive PD-L1 Assays

GS16. *Expert opinion*—Pathologists issuing reports for PD-L1 testing should successfully undergo training that is (i) endorsed by the manufacturer of an approved PD-L1 IHC Kit, or (ii) endorsed by a professional organization (eg, Canadian Association of Pathologists, International Quality Network for Pathology, etc.), or (iii) provided by a pathologist who has undergone training as indicated in (i) or (ii) and in a train-the-trainer format. The optimal goal for such training should be to achieve 90% sensitivity and 90% specificity of the previously validated readout.^{131–136}

Explanatory note: Readout accuracy (sensitivity and specificity) needs to be demonstrated for relevant cutoff(s), rather than for overall agreement with a designated reference standard or concordance. Documentation and retention of readout validation evidence apply as per relevant documentation and retention of evidence of any IHC assay validation.

GS17. *Expert opinion*—Reporting of computer-assisted PD-L1 readout (image analysis) for the purpose of patient selection for immunotherapy is not recommended at this time.^{137–142}

Explanatory note: Some PD-L1 readout methods may be more amenable to image analysis than others; therefore, any recommendations at this time may not be universally applicable for all different types of readouts. At this time, there is insufficient evidence to recommend the use of image analysis for the readout of PD-L1 assays with confidence. Pathologist-assisted image analysis may prove to be a valuable tool as it further develops.

GS18. *Recommendation*—A structured (synoptic) format should be used for PD-L1 reporting. Searchable formats are recommended to facilitate audit of institutional or pathologist-specific prevalence.^{131,143–147}

No Explanatory note.

GS19. *Recommendation*—In the comments section of the report, the following information should be included: (i) the drug and disease for which the assay was run, and (ii) the assay that was performed, and on what platform, (iii) the clinically relevant cutoff point for the 3D combination, (iv) the readout parameter being reported (eg, Tumour Proportion Score, Combined Positive Score, Immune Cells), and (v) the clinically relevant cutoff category within which the result falls.^{148–157}

No Explanatory note.

GS20. *Expert opinion*—If the specimen was inadequate for testing, this may be stated in the report along with the reason why the specimen was inadequate.

No Explanatory note.

GS21. *Expert opinion*—Results with cytology samples may be reported in the same manner recommended for solid tissue samples (see GS19).

No Explanatory note.

GS22. *Expert opinion*—Results with fresh-cut sections from archived paraffin blocks or previously-cut and stored unstained sections that are older than the PD-L1 IHC Kit manufacturer's recommendation may be reported only if positive and it is the pathologist's assessment that, based on presented evidence, the result is not false-positive (eg, pattern of staining being membranous, appropriate internal and external controls are present and show expected range of positive and negative result). "Positive result" is here defined as a positive result for a specific purpose and a specific patient context (eg, the staining of inflammatory cells has no relevance for some PD-L1 assays while for others, they are the principal cell type assessed to determine whether the sample is to be designated as being a "positive" result). Therefore, some positivity identified in a lesion does not qualify the results as reportable. Negative results may be reported as "no result" and potential for false-negative results with older samples may be included in the comment section.^{158,159}

No Explanatory note.

Recommendations for Quality Assurance

Validation

GS23. *Strong recommendation*—Validation of PD-L1 IHC predictive assays must be fit-for-purpose in accordance with the 3D (Drug-Disease-Diagnostic assay) approach.^{58,65,76,77,94}

Explanatory note: This statement mirrors GS01 but from the perspective of the laboratory. The Drug-Disease combination informs the laboratory of the PD-L1 IHC Kit that can either (i) be verified and run as the companion/complementary Diagnostic assay for the chosen Drug-Disease combination, or (ii) be used as the reference standard assay if the laboratory decides to develop, validate, run, and maintain an LDT for the chosen Drug-Disease combination. The Disease component of the 3D approach also informs the laboratory of the type of validation cases necessary to collect in order to create the tissue tools that will allow for ongoing monitoring of analytical sensitivity and specificity on a daily basis and of diagnostic sensitivity and specificity on a periodic basis. Also, see Explanatory note GS01.

GS24. *Strong recommendation*—When a laboratory decides that an LDT will be used instead of a regulatory agency-approved IHC kit for the same purpose, the IHC laboratory must provide evidence of successful methodology transfer. This must include (i) technical validation and (ii) indirect clinical validation by using the relevant reference standard where not already published or established (ie, the regulatory agency-approved kit for the same purpose).^{58,59,65,77–79,86,91}

No Explanatory note.

GS25. *Recommendation*—Indirect clinical validation can be performed when there is a recognized reference standard or gold standard that has been validated in a published clinical trial for that specific purpose. Samples for indirect clinical validation should include at least 50

randomly selected, clinically relevant positive cases, and at least 50 negative cases identified in the screening process.⁵⁵

Explanatory note: The tissues for indirect clinical validation may originate from the institution that is performing the indirect clinical validation for a newly developed LDT, but it also may originate from a PT program's reference laboratory as well as from multi-institutional sources. See Terminology section for more information about indirect clinical validation and explanatory notes for GS05 and GS26.

GS26. Strong recommendation—PD-L1 IHC LDTs that have been clinically validated or indirectly clinically validated in published literature must still be technically validated by the clinical IHC laboratory before being put in use as a PD-L1 predictive biomarker assay.^{58,59,65,86}

Explanatory note: A properly validated LDT in the context of PD-L1 being a patient selection biomarker in immunoncology is one that has successfully undergone both clinical validation (direct or indirect) and technical validation. Clinical validation may be direct (where the reference standard result is based on the clinical responses of clinical trial participants) or indirect (where the reference standard result is based on the result generated by the companion diagnostic assay that was developed from clinical responses of clinical trial participants). Evidence for clinical validation (direct or indirect) of an LDT may be generated by the laboratory or alternatively, if available, may be derived from the literature. Clinical validation (direct or indirect) evaluates the total test. Evidence for technical validation must be generated by the laboratory—such evidence cannot be derived from literature. Technical validation is performed on the analytical phase of the total test, namely the protocol and the readout. Therefore, direct clinical validation is based on clinical outcomes (the study of patients), indirect clinical validation on the reference standard assay results (the study of cases), and technical validation on the protocol and/or readout results with validation samples (ie, QA/QC tissue tools).

GS27. Strong recommendation—Health Canada-approved, FDA-approved, and CE-marked PD-L1 IHC Kits are already directly or indirectly clinically validated in clinical trials; however, their analytical performance must be verified according to manufacturer's instructions before use. If no specific instructions are provided by the manufacturer, general recommendations for Tier 1 technical validation apply (analytical sensitivity, analytical specificity, and reproducibility) by using iCAPCs or iCAPCs-like tissues/calibrators with at least 3 successful, independent verification runs.^{58,160}

No Explanatory note.

GS28. Strong recommendation—As per recommendations above, if multiplex assays (eg, double IHC staining for PD-L1 and TTF-1) are to be used, they must have undergone technical validation. The possible impact of multiplexing on diagnostic accuracy of each biomarker must be assessed against single IHC staining using a Health Canada-approved, FDA-approved, or CE-marked IHC kit or other fit-for-purpose designated reference standard.^{58,65,94}

No Explanatory note.

GS29. Strong recommendation—The PD-L1 IHC readout must be validated for diagnostic accuracy (readout accuracy) by each pathologist reporting PD-L1 IHC results as a categorical variable; in addition, validation of readout precision must also be conducted when applicable (eg, if reporting exact % positive cells as a continuous variable). This applies for both regulatory agency-approved IHC kits and LDTs.^{58,66,76,160}

Explanatory note: There are online and in-person resources available for readout training. In addition, pathologists may avail themselves of EQA PT tools developed specifically for pathologist readout proficiency where and when available.

GS30. Strong recommendation—Laboratories using regulatory agency-approved PD-L1 IHC Kits in a manner that deviates from the manufacturer's instructions must validate the deviation.^{58,65,86,161}

Explanatory note: Deviations may occur in the pre-analytical phase, the analytical phase, or the post-analytical phase of the total test. See also GS07.

Quality Control

GS31. Strong recommendation—Laboratories must run external on-slide controls for every slide.^{58,64,66,162,163}

No Explanatory note.

GS32. Strong recommendation—The minimum composition of external controls for FDA (or other regulatory body)-approved PD-L1 IHC Kits must be in compliance with the Kit manufacturer's instructions.^{148,155,156,161,164}

No Explanatory note.

GS33. Expert opinion—The minimum composition of external on-slide controls for LDTs ideally will mirror the requirements of the corresponding fit-for-purpose FDA (or other regulatory body)-approved kit.

No Explanatory note.

GS34. Expert opinion—Where not already included as part of the required minimum composition, external on-slide controls for PD-L1 testing may include the following elements: (i) lesional tissue showing positivity for PD-L1 in a specified cell population, (ii) lesional tissue showing negativity for PD-L1 in a specified cell population, (iii) low expressor tissue in accordance with iCAPCs-like principles (eg, macrophages in germinal centres of tonsil).^{58,64,163}

No Explanatory note.

GS35. Recommendation—The laboratory should assess the external on-slide control by light microscopy prior to a slide being released from the laboratory. When external on-slide controls demonstrate either unacceptable background or lack of signal in any positive control tissue, such IHC slides should be failed by the laboratory.^{64,162}

No Explanatory note.

EQA—PT

GS36. Recommendation—All laboratories performing PD-L1 testing for patient selection should participate in fit-for-purpose PT at least every six months where available and required by relevant regulatory bodies.^{165–169}

Explanatory note: A fit-for-purpose PT challenge follows the 3D approach. The Drug-Disease combination determines the Diagnostic assay that will be used to generate the reference standard results against which the results of the participating laboratories are measured. The Disease determines the selection of the tumor type and the readout that will be used in the PT run to calculate diagnostic accuracy.

GS37. Recommendation—Laboratories should select a PT program that ideally informs participating laboratories on the accuracy of the PD-L1 IHC protocol and the pathologists' readout as determined by central expert assessment.^{165,170,171} *If no such program is available/accessibile, inter-laboratory exchange of samples and information with a provincial or national reference laboratory should be performed.*^{172,173}

No Explanatory note.

GS38. Strong recommendation—Participation in PT, even when PT is fit-for-purpose, is not acceptable as a substitute for clinical and technical validation.^{58,59,65,86}

Explanatory note: When fit-for-purpose and when designed appropriately, PT may be a substitute for indirect clinical validation (but not for clinical validation or technical validation).

CONCLUSIONS

In situ biomarker testing is not defined solely by the detection of biological gene products in human tissue sections; rather, such testing must always be accompanied by the appropriate medical context to be meaningful for patient care. For diagnostic biomarker testing, the medical context may be the impact of tissue specificity (eg, tumor type) on the meaningfulness of test results, whereas, for predictive biomarker testing, the medical context typically also includes a specific therapeutic agent in addition to tumor type or tissue specificity. Therefore, the current reality of PD-L1 testing in immuno-oncology is such that detection of PD-L1 protein expression is only meaningful in the context of the tumor(s)/tissue type(s) being tested, for potential response to a particular therapeutic agent based on data generated by clinical trials. Given the evolving landscape for PD-L1 testing, the intention of the CAP-ACP National Standards Committee for High Complexity Testing is to periodically update these GSs as long as PD-L1 remains a relevant biomarker for patient selection in immuno-oncology.

REFERENCES

- Dong H, Zhu G, Tamada K, et al. B7-H1, a third member of the B7 family, co-stimulates T-cell proliferation and interleukin-10 secretion. *Nat Med*. 1999;5:1365–1369.
- Freeman GJ, Long AJ, Iwai Y, et al. Engagement of the PD-1 immunoinhibitory receptor by a novel B7 family member leads to negative regulation of lymphocyte activation. *J Exp Med*. 2000;192:1027–1034.
- Latchman YE, Liang SC, Wu Y, et al. PD-L1-deficient mice show that PD-L1 on T cells, antigen-presenting cells, and host tissues negatively regulates T cells. *Proc Natl Acad Sci USA*. 2004;101:10691–10696.
- Dong H, Strome SE, Salomao DR, et al. Tumor-associated B7-H1 promotes T-cell apoptosis: a potential mechanism of immune evasion. *Nat Med*. 2002;8:793–800.
- Inaguma S, Wang Z, Lasota J, et al. Comprehensive immunohistochemical study of programmed cell death ligand 1 (PD-L1): analysis in 5536 cases revealed consistent expression in trophoblastic tumors. *Am J Surg Pathol*. 2016;40:1133–1142.
- Veras E, Kurman RJ, Wang T-L, et al. PD-L1 expression in human placentas and gestational trophoblastic diseases. *Int J Gynecol Pathol*. 2017;36:146–153.
- Ancevski Hunter K, Socinski MA, Villaruz LC. PD-L1 testing in guiding patient selection for PD-1/PD-11 inhibitor therapy in lung cancer. *Mol Diagn Ther*. 2018;22:1–10.
- Brody R, Zhang Y, Ballas M, et al. PD-L1 expression in advanced NSCLC: insights into risk stratification and treatment selection from a systematic literature review. *Lung Cancer*. 2017;112:200–215.
- Xiang X, Yu P-C, Long D, et al. Prognostic value of PD-L1 expression in patients with primary solid tumors. *Oncotarget*. 2018;9:5058–5072.
- Aguar PN, De Mello RA, Hall P, et al. PD-L1 expression as a predictive biomarker in advanced non-small-cell lung cancer: updated survival data. *Immunotherapy*. 2017;9:499–506.
- 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;31:1381–1390.
- 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;13:377–386.
- Hersom M, Jørgensen JT. Companion and complementary diagnostics-focus on PD-L1 expression assays for PD-1/PD-L1 checkpoint inhibitors in non-small cell lung cancer. *Ther Drug Monit*. 2018;40:9–16.
- Tibaldi C, Lunghi A, Baldini E. Use of programmed cell death protein ligand 1 assay to predict the outcomes of non-small cell lung cancer patients treated with immune checkpoint inhibitors. *World J Clin Oncol*. 2017;8:320–328.
- Guan J, Lim KS, Mekhail T, et al. Programmed death ligand-1 (PD-L1) expression in the programmed death receptor-1 (PD-1)/PD-L1 blockade: a key player against various cancers. *Arch Pathol Lab Med*. 2017;141:851–861.
- Kogashiwa Y, Yasuda M, Sakurai H, et al. PD-L1 expression confers better prognosis in locally advanced oral squamous cell carcinoma. *Anticancer Res*. 2017;37:1417–1424.
- Hodi FS, O'Day SJ, McDermott DF, et al. Improved survival with ipilimumab in patients with metastatic melanoma. *N Engl J Med*. 2010;363:711–723.
- Garon EB, Rizvi NA, Hui R, et al. Pembrolizumab for the treatment of non-small-cell lung cancer. *N Engl J Med*. 2015;372:2018–2028.
- Herbst RS, Baas P, Kim D-W, 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;387:1540–1550.
- Reck M, Rodríguez-Abreu D, Robinson AG, et al. Pembrolizumab versus chemotherapy for PD-L1-positive non-small-cell lung cancer. *N Engl J Med*. 2016;375:1823–1833.
- Khazin S, Abernethy AP, Nussbaum NC, et al. Characteristics of real-world metastatic non-small cell lung cancer patients treated with nivolumab and pembrolizumab during the year following approval. *Oncologist*. 2018;23:328–336.
- Gong J, Chehraz-Raffle A, Reddi S, et al. Development of PD-1 and PD-L1 inhibitors as a form of cancer immunotherapy: a comprehensive review of registration trials and future considerations. *J Immunother Cancer*. 2018;6:8.
- Fuchs CS, Doi T, Jang RW, et al. Safety and efficacy of pembrolizumab monotherapy in patients with previously treated advanced gastric and gastroesophageal junction cancer: phase 2 Clinical KEYNOTE-059 Trial. *JAMA Oncol*. 2018;4:e180013.
- Schellens JHM, Marabelle A, Zeigenfuss S, et al. Pembrolizumab for previously treated advanced cervical squamous cell cancer:

- Preliminary results from the phase 2 KEYNOTE-158 study. *J Clin Oncol*. 2017;35:5514–5514.
25. Suzman DL, Agrawal S, Ning Y-M, et al. FDA approval summary: atezolizumab or pembrolizumab for the treatment of patients with advanced urothelial carcinoma ineligible for cisplatin-containing chemotherapy. *Oncologist*. 2019;24:563–569.
 26. Schmid P, Adams S, Rugo HS, et al. Atezolizumab and nab-paclitaxel in advanced triple-negative breast cancer. *N Engl J Med*. 2018;379:2108–2121.
 27. Siu LL, Even C, Mesía R, et al. Safety and efficacy of durvalumab with or without tremelimumab in patients with PD-L1-low/negative recurrent or metastatic HNSCC: the Phase 2 CONDOR Randomized Clinical Trial. *JAMA Oncol*. 2019;5:195–203.
 28. Ott PA, Bang Y-J, Piha-Paul SA, et al. T-cell-inflamed gene-expression profile, programmed death ligand 1 expression, and tumor mutational burden predict efficacy in patients treated with pembrolizumab across 20 cancers: KEYNOTE-028. *J Clin Oncol*. 2019;37:318–327.
 29. Sharma P, Retz M, Siefker-Radtke A, et al. Nivolumab in metastatic urothelial carcinoma after platinum therapy (CheckMate 275): a multicentre, single-arm, phase 2 trial. *Lancet Oncol*. 2017;18:312–322.
 30. 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;373:1627–1639.
 31. Ferris RL, Blumenschein G, Fayette J, et al. Nivolumab for recurrent squamous-cell carcinoma of the head and neck. *N Engl J Med*. 2016;375:1856–1867.
 32. Larkin J, Chiarion-Sileni V, Gonzalez R, et al. Combined nivolumab and ipilimumab or monotherapy in untreated melanoma. *N Engl J Med*. 2015;373:23–34.
 33. 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;387:1837–1846.
 34. 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;389:255–265.
 35. Roach C, Zhang N, Corigliano E, et al. Development of a companion diagnostic PD-L1 immunohistochemistry assay for pembrolizumab therapy in non-small-cell lung cancer. *Appl Immunohistochem Mol Morphol*. 2016;24:392–397.
 36. Dolled-Filhart M, Roach C, Toland G, et al. Development of a companion diagnostic for pembrolizumab in non-small cell lung cancer using immunohistochemistry for programmed death ligand-1. *Arch Pathol Lab Med*. 2016;140:1243–1249.
 37. Phillips T, Millett MM, Zhang X, et al. Development of a diagnostic programmed cell death 1-ligand 1 immunohistochemistry assay for nivolumab therapy in melanoma. *Appl Immunohistochem Mol Morphol*. 2018;26:6–12.
 38. Phillips T, Simmons P, Inzunza HD, et al. Development of an automated PD-L1 immunohistochemistry (IHC) assay for non-small cell lung cancer. *Appl Immunohistochem Mol Morphol*. 2015;23:541–549.
 39. Vennapusa B, Baker B, Kowanz M, et al. Development of a PD-L1 complementary diagnostic immunohistochemistry assay (SP142) for atezolizumab. *Appl Immunohistochem Mol Morphol*. 2018;27:92–100.
 40. Rebelatto MC, Midha A, Mistry A, et al. Development of a programmed cell death ligand-1 immunohistochemical assay validated for analysis of non-small cell lung cancer and head and neck squamous cell carcinoma. *Diagn Pathol*. 2016;11:95.
 41. Zajac M, Boothman A-M, Ben Y, et al. Analytical validation and clinical utility of an immunohistochemical programmed death ligand-1 diagnostic assay and combined tumor and immune cell scoring algorithm for durvalumab in urothelial carcinoma. *Arch Pathol Lab Med*. 2018;143:722–731.
 42. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ*. 2015;351:h5527.
 43. Schünemann HJ. Interpreting GRADE's levels of certainty or quality of the evidence: GRADE for statisticians, considering review information size or less emphasis on imprecision? *J Clin Epidemiol*. 2016;75:6–15.
 44. Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155:529–536.
 45. Shiffman RN, Shekelle P, Overhage JM, et al. Standardized reporting of clinical practice guidelines: a proposal from the Conference on Guideline Standardization. *Ann Intern Med*. 2003;139:493–498.
 46. Brouwers MC, Kerkvliet K, Spithoff K, et al. The AGREE Reporting Checklist: a tool to improve reporting of clinical practice guidelines. *BMJ*. 2016;352:i1152.
 47. Balshem H, Helfand M, Schünemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol*. 2011;64:401–406.
 48. Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol*. 2011;64:383–394.
 49. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 2. Framing the question and deciding on important outcomes. *J Clin Epidemiol*. 2011;64:395–400.
 50. Brozek JL, Akl EA, Jaeschke R, et al. Grading quality of evidence and strength of recommendations in clinical practice guidelines: part 2 of 3. The GRADE approach to grading quality of evidence about diagnostic tests and strategies. *Allergy*. 2009;64:1109–1116.
 51. Brozek JL, Akl EA, Compalati E, et al. Grading quality of evidence and strength of recommendations in clinical practice guidelines: part 3 of 3. The GRADE approach to developing recommendations. *Allergy*. 2011;66:588–595.
 52. Chen Y, Yang K, Marušić A, et al. A reporting tool for practice guidelines in health care: the RIGHT statement. *Ann Intern Med*. 2017;166:128–132.
 53. Jones CM, Ashrafian H, Darzi A, et al. Guidelines for diagnostic tests and diagnostic accuracy in surgical research. *J Invest Surg*. 2010;23:57–65.
 54. Straus SE, Shepperd S. Challenges in guideline methodology. *J Clin Epidemiol*. 2011;64:347–348.
 55. Garrett PE, Lasky FD, Meier KL, et al. *User Protocol for Evaluation of Qualitative Test Performance: Approved Guideline*. Wayne, PA: Clinical and Laboratory Standards Institute; 2008.
 56. Joint Committee for Guides in Metrology. *International Vocabulary of Metrology—Basic and General Concepts and Associated Terms (VIM)*, 3rd ed. 2012.
 57. Cohen JF, Korevaar DA, Altman DG, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open*. 2016;6:e012799.
 58. Torlakovic EE, Cheung CC, D'Arrigo C, et al. Evolution of quality assurance for clinical immunohistochemistry in the era of precision medicine. Part 3: technical validation of immunohistochemistry (IHC) assays in clinical IHC laboratories. *Appl Immunohistochem Mol Morphol*. 2017;25:151–159.
 59. Masucci GV, Cesano A, Hawtin R, et al. Validation of biomarkers to predict response to immunotherapy in cancer: volume I—pre-analytical and analytical validation. *J Immunother Cancer*. 2016;4:76.
 60. Bossuyt PMM, Reitsma JB, Linnet K, et al. Beyond diagnostic accuracy: the clinical utility of diagnostic tests. *Clin Chem*. 2012;58:1636–1643.
 61. US Food and Drug Administration. In vitro diagnostics—companion diagnostics. Available at: www.fda.gov/medicaldevices/productsandmedicalprocedures/invitrodiagnostics/ucm407297.htm. Accessed April 7, 2019.
 62. Scheerens H, Malong A, Bassett K, et al. Current status of companion and complementary diagnostics: strategic considerations for development and launch: options for companion and complementary diagnostics. *Clin Transl Sci*. 2017;10:84–92.
 63. US Food and Drug Administration. *FDA-NIH Biomarker Working Group—Diagnostic Biomarker*. Silver Spring, MD: Food and Drug Administration; 2016.

64. Torlakovic EE, Nielsen S, Francis G, et al. Standardization of positive controls in diagnostic immunohistochemistry: recommendations from the International Ad Hoc Expert Committee. *Appl Immunohistochem Mol Morphol*. 2015;23:1–18.
65. 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;138:1432–1443.
66. Torlakovic EE, Cheung CC, D'Arrigo C, et al. Evolution of quality assurance for clinical immunohistochemistry in the era of precision medicine—part 2: immunohistochemistry test performance characteristics. *Appl Immunohistochem Mol Morphol*. 2017;25:79–85.
67. Canadian Standards Association. Requirements for the design, development, and validation of laboratory-developed tests; 2018. Available at: https://store.csagroup.org/ccrz_ProductDetails?viewState=DetailView&cartID=&sku=Z316.8-18&isCSRFlow=true&portalUser=&store=&cccl=en_US. Accessed September 2, 2018.
68. US Food and Drug Administration. *FDA-NIH Biomarker Working Group—BEST (Biomarkers, EndpointS, and other Tools) Resource Understanding Prognostic versus Predictive Biomarkers*. Silver Spring; Bethesda, MD: Food and Drug Administration National Institutes of Health (US); 2016:6.
69. Lee JW, Figeys D, Vasilescu J. Biomarker assay translation from discovery to clinical studies in cancer drug development: quantification of emerging protein biomarkers. *Adv Cancer Res*. 2007;96:269–298.
70. Lee JW, Weiner RS, Sailstad JM, et al. Method validation and measurement of biomarkers in nonclinical and clinical samples in drug development: a conference report. *Pharm Res*. 2005;22:499–511.
71. Corsello SM, Bittker JA, Liu Z, et al. The Drug Repurposing Hub: a next-generation drug library and information resource. *Nat Med*. 2017;23:405–408.
72. Pushpakom S, Iorio F, Eyers PA, et al. Drug repurposing: progress, challenges and recommendations. *Nat Rev Drug Discov*. 2019;18:41–58.
73. Kulangara K, Zhang N, Corigliano E, et al. Clinical utility of the combined positive score for programmed death ligand-1 expression and the approval of pembrolizumab for treatment of gastric cancer. *Arch Pathol Lab Med*. 2019;143:330–337.
74. Marchiò C, Dowsett M, Reis-Filho JS. Revisiting the technical validation of tumour biomarker assays: how to open a Pandora's box. *BMC Med*. 2011;9:41.
75. ISO. Quality management systems—fundamentals and vocabulary. Available at: www.iso.org/obp/ui/#iso:std:iso:9000:ed-4:v1:en. Accessed April 7, 2019.
76. Cheung CC, D'Arrigo C, Dietel M, et al. Evolution of quality assurance for clinical immunohistochemistry in the era of precision medicine: part 1: fit-for-purpose approach to classification of clinical immunohistochemistry biomarkers. *Appl Immunohistochem Mol Morphol*. 2017;25:4–11.
77. Lee JW, Devanarayan V, Barrett YC, et al. Fit-for-purpose method development and validation for successful biomarker measurement. *Pharm Res*. 2006;23:312–328.
78. Cummings J, Raynaud F, Jones L, et al. Fit-for-purpose biomarker method validation for application in clinical trials of anticancer drugs. *Br J Cancer*. 2010;103:1313–1317.
79. Marrero A, Lawrence S, Wilsker D, et al. Translating pharmacodynamic biomarkers from bench to bedside: analytical validation and fit-for-purpose studies to qualify multiplex immunofluorescent assays for use on clinical core biopsy specimens. *Semin Oncol*. 2016;43:453–463.
80. Rosenberg JE, Hoffman-Censits J, Powles T, et al. Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial. *Lancet*. 2016;387:1909–1920.
81. Powles T, O'Donnell PH, Massard C, et al. Efficacy and safety of durvalumab in locally advanced or metastatic urothelial carcinoma: updated results from a phase 1/2 open-label study. *JAMA Oncol*. 2017;3:e172411.
82. Massard C, Gordon MS, Sharma S, et al. Safety and efficacy of durvalumab (MEDI4736), an anti-programmed cell death ligand-1 immune checkpoint inhibitor, in patients with advanced urothelial bladder cancer. *J Clin Oncol*. 2016;34:3119–3125.
83. Herbst RS, Soria J-C, Kowanetz M, et al. Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. *Nature*. 2014;515:563–567.
84. Bellmunt J, Wit R, Vaughn DJ, et al. Pembrolizumab as second-line therapy for advanced urothelial carcinoma. *N Engl J Med*. 2017;376:1015–1026.
85. Balar AV, Castellano D, O'Donnell PH, et al. First-line pembrolizumab in cisplatin-ineligible patients with locally advanced and unresectable or metastatic urothelial cancer (KEYNOTE-052): a multicentre, single-arm, phase 2 study. *Lancet Oncol*. 2017;18:1483–1492.
86. Dobbin KK, Cesano A, Alvarez J, et al. Validation of biomarkers to predict response to immunotherapy in cancer: Volume II — clinical validation and regulatory considerations. *J Immunother Cancer*. 2016;4:77.
87. Khleif SN, Doroshow JH, Hait WN, et al. AACR-FDA-NCI cancer biomarkers collaborative consensus report: advancing the use of biomarkers in cancer drug development. *Clin Cancer Res*. 2010;16:3299–3318.
88. Woodcock J, Woosley R. The FDA critical path initiative and its influence on new drug development*. *Annu Rev Med*. 2008;59:1–12.
89. US Food and Drug Administration. List of cleared or approved companion diagnostic device; 2017. Available at: www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/InVitroDiagnostics/ucm301431.htm. Accessed January 17, 2019.
90. Torlakovic E, Lim HJ, Adam J, et al. Interchangeability of PD-L1 immunohistochemistry assays: meta-analysis of diagnostic accuracy. *Mod Pathol*. 2019. [Epub ahead of print].
91. National Biomarker Development Alliance (NBDA). Biomarker qualification; 2018.
92. US Food and Drug Administration. Guidance documents (Medical Devices and Radiation-Emitting Products)—Guidance for Submission of Immunohistochemistry Applications to the FDA; Final Guidance for Industry; 2016. Available at: www.fda.gov/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm094002.htm. Accessed July 24, 2016.
93. Cheung CC, Lim HJ, Garratt J, et al. Diagnostic accuracy in fit-for-purpose PD-L1 testing. *Appl Immunohistochem Mol Morphol*. 2019;27:251–257.
94. Elliott K, McQuaid S, Salto-Tellez M, et al. Immunohistochemistry should undergo robust validation equivalent to that of molecular diagnostics. *J Clin Pathol*. 2015;68:766–770.
95. Taylor CR. FDA issues final rule for classification and reclassification of immunochemistry reagents and kits. *Am J Clin Pathol*. 1999;111:443–444.
96. Engel KB, Moore HM. Effects of preanalytical variables on the detection of proteins by immunohistochemistry in formalin-fixed, paraffin-embedded tissue. *Arch Pathol Lab Med*. 2011;135:537–543.
97. 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;30:340–349.
98. McLaughlin J, Han G, Schalper KA, et al. Quantitative assessment of the heterogeneity of PD-L1 expression in non-small-cell lung cancer. *JAMA Oncol*. 2016;2:46–54.
99. Scorer P, Scott M, Lawson N, et al. Consistency of tumor and immune cell programmed cell death ligand-1 expression within and between tumor blocks using the VENTANA SP263 assay. *Diagn Pathol*. 2018;13:47.
100. López JI, Pulido R, Cortés JM, et al. Potential impact of PD-L1 (SP-142) immunohistochemical heterogeneity in clear cell renal cell carcinoma immunotherapy. *Pathol Res Pract*. 2018;214:1110–1114.
101. López JI, Pulido R, Lawrie CH, et al. Loss of PD-L1 (SP-142) expression characterizes renal vein tumor thrombus microenvironment in clear cell renal cell carcinoma. *Ann Diagn Pathol*. 2018;34:89–93.
102. Madore J, Vilain RE, Menzies AM, et al. PD-L1 expression in melanoma shows marked heterogeneity within and between

- patients: implications for anti-PD-1/PD-L1 clinical trials. *Pigment Cell Melanoma Res.* 2015;28:245–253.
103. Callea M, Albiges L, Gupta M, et al. Differential expression of PD-L1 between primary and metastatic sites in clear-cell renal cell carcinoma. *Cancer Immunol Res.* 2015;3:1158–1164.
 104. 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;12:458–466.
 105. 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.
 106. Liu Y, Dong Z, Jiang T, et al. Heterogeneity of PD-L1 expression among the different histological components and metastatic lymph nodes in patients with resected lung adenocarcinoma. *Clin Lung Cancer.* 2018;19:e421–e430.
 107. Dill EA, Gru AA, Atkins KA, et al. PD-L1 expression and intratumoral heterogeneity across breast cancer subtypes and stages: an assessment of 245 primary and 40 metastatic tumors. *Am J Surg Pathol.* 2017;41:334–342.
 108. 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;25:453–459.
 109. Casadevall D, Clavé S, Taus Á, et al. Heterogeneity of tumor and immune cell PD-L1 expression and lymphocyte counts in surgical NSCLC samples. *Clin Lung Cancer.* 2017;18:682.e5–691.e5.
 110. Li C, Huang C, Mok TS, et al. Comparison of 22C3 PD-L1 expression between surgically resected specimens and paired tissue microarrays in non-small cell lung cancer. *J Thorac Oncol.* 2017;12:1536–1543.
 111. Cho JH, Sorensen SF, Choi Y-L, et al. Programmed death ligand 1 expression in paired non-small cell lung cancer tumor samples. *Clin Lung Cancer.* 2017;18:e473–e479.
 112. Kim S, Kim M-Y, Koh J, et al. Programmed death-1 ligand 1 and 2 are highly expressed in pleomorphic carcinomas of the lung: comparison of sarcomatous and carcinomatous areas. *Eur J Cancer.* 2015;51:2698–2707.
 113. Liu T, Greenberg M, Wentland C, et al. PD-L1 expression and CD8+ infiltration shows heterogeneity in juvenile recurrent respiratory papillomatosis. *Int J Pediatr Otorhinolaryngol.* 2017;95:133–138.
 114. Wang HB, Yao H, Li CS, et al. Rise of PD-L1 expression during metastasis of colorectal cancer: implications for immunotherapy: rise of PD-L1 during metastasis. *J Dig Dis.* 2017;18:574–581.
 115. El Jabbour T, Ross JS, Sheehan CE, et al. PD-L1 protein expression in tumour cells and immune cells in mismatch repair protein-deficient and -proficient colorectal cancer: the foundation study using the SP142 antibody and whole section immunohistochemistry. *J Clin Pathol.* 2018;71:46–51.
 116. Nakamura S, Hayashi K, Imaoka Y, et al. Intratumoral heterogeneity of programmed cell death ligand-1 expression is common in lung cancer. *PLoS One.* 2017;12:e0186192.
 117. Li M, Li A, Zhou S, et al. Heterogeneity of PD-L1 expression in primary tumors and paired lymph node metastases of triple negative breast cancer. *BMC Cancer.* 2018;18:4.
 118. Zhou J, Gong Z, Jia Q, et al. Programmed death ligand 1 expression and CD8+ tumor-infiltrating lymphocyte density differences between paired primary and brain metastatic lesions in non-small cell lung cancer. *Biochem Biophys Res Commun.* 2018;498:751–757.
 119. Minnema-Luiting J, Vroman H, Aerts J, et al. Heterogeneity in immune cell content in malignant pleural mesothelioma. *Int J Mol Sci.* 2018;19:1041.
 120. Castello A, Grizzi F, Toschi L, et al. Tumor heterogeneity, hypoxia, and immune markers in surgically resected non-small-cell lung cancer. *Nucl Med Commun.* 2018;39:636–644.
 121. Mansfield AS, Murphy SJ, Peikert T, et al. Heterogeneity of programmed cell death ligand 1 expression in multifocal lung cancer. *Clin Cancer Res.* 2016;22:2177–2182.
 122. Mansfield AS, Aubry MC, Moser JC, et al. Temporal and spatial discordance of programmed cell death-ligand 1 expression and lymphocyte tumor infiltration between paired primary lesions and brain metastases in lung cancer. *Ann Oncol.* 2016;27:1953–1958.
 123. Wang H, Agulnik J, Kasymjanova G, et al. Cytology cell blocks are suitable for immunohistochemical testing for PD-L1 in lung cancer. *Ann Oncol.* 2018;29:1417–1422.
 124. Torous VF, Rangachari D, Gallant BP, et al. 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;7:133–141.
 125. Stoy SP, Rosen L, Mueller J, et al. Programmed death-ligand 1 testing of lung cancer cytology specimens obtained with bronchoscopy: PD-L1 testing in lung cancer cytology. *Cancer Cytopathol.* 2018;126:122–128.
 126. Sakakibara R, Inamura K, Tambo Y, et al. EBUS-TBNA as a promising method for the evaluation of tumor PD-L1 expression in lung cancer. *Clin Lung Cancer.* 2017;18:527.e1–534.e1.
 127. Noll B, Wang W-L, Gong Y, et al. Programmed death ligand 1 testing in non-small cell lung carcinoma cytology cell block and aspirate smear preparations: PD-L1 testing in lung cytology. *Cancer Cytopathol.* 2018;126:342–352.
 128. Ilie M, Juco J, Huang L, et al. 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: PD-L1 22C3 cytology-based LDTs in NSCLC. *Cancer Cytopathol.* 2018;126:264–274.
 129. Heymann JJ, Bulman WA, Swinarski D, et al. PD-L1 expression in non-small cell lung carcinoma: comparison among cytology, small biopsy, and surgical resection specimens. *Cancer Cytopathol.* 2017;125:896–907.
 130. Biswas A, Leon ME, Drew P, et al. Clinical performance of endobronchial ultrasound-guided transbronchial needle aspiration for assessing programmed death ligand-1 expression in nonsmall cell lung cancer. *Diagn Cytopathol.* 2018;46:378–383.
 131. Cagle PT, Sholl LM, Lindeman NI, et al. Template for reporting results of biomarker testing of specimens from patients with non-small cell carcinoma of the lung. *Arch Pathol Lab Med.* 2014;138:171–174.
 132. 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;23:4569–4577.
 133. Brunström 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 Pathol.* 2017;30:1411–1421.
 134. 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;3:1051.
 135. Scheel AH, Dietel M, Heukamp LC, et al. Harmonized PD-L1 immunohistochemistry for pulmonary squamous-cell and adenocarcinomas. *Mod Pathol.* 2016;29:1165–1172.
 136. Scheel AH, Baenfer G, Baretton G, et al. Interlaboratory concordance of PD-L1 immunohistochemistry for non-small-cell lung cancer. *Histopathology.* 2018;72:449–459.
 137. Barnes M, Bai I, Nguyen K, et al. P2.01-043 pathologist agreement rates of PD-L1 tumor and immune cell quantitation using digital read, field-of-view, and whole tumor image analysis: topic: immune mechanisms in thoracic cancer and targeted therapy. *J Thorac Oncol.* 2017;12:S811–S812.
 138. Koelzer VH, Gisler A, Hanhart JC, et al. Digital image analysis improves precision of PD-L1 scoring in cutaneous melanoma. *Histopathology.* 2018;73:397–406.
 139. Wiestler T, Widmaier M, Walker J, et al. 103PComparison of continuous measures across diagnostic PD-L1 assays using image analysis. *Ann Oncol.* 2017;28(suppl 5).
 140. Parra ER, Behrens C, Rodriguez-Canales J, et al. Image analysis-based assessment of PD-L1 and tumor-associated immune cells density supports distinct intratumoral microenvironment groups in non-small cell lung carcinoma patients. *Clin Cancer Res.* 2016;22:6278–6289.

141. Kapil A, Meier A, Zuraw A, et al. Deep semi supervised generative learning for automated PD-L1 tumor cell scoring on NSCLC tissue needle biopsies. *Sci Rep*. 2018;8:17343.
142. Parra ER, Villalobos P, Mino B, et al. 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;26:83.
143. Renshaw SA, Mena-Allauca M, Touriz M, et al. The impact of template format on the completeness of surgical pathology reports. *Arch Pathol Lab Med*. 2014;138:121–124.
144. Renshaw AA, Mena-Allauca M, Gould EW, et al. Synoptic reporting: evidence-based review and future directions. *JCO Clin Cancer Inform*. 2018;2:1–9.
145. Newkirk M, Pamplin JC, Kuwamoto R, et al. Checklists change communication about key elements of patient care. *J Trauma Acute Care Surg*. 2012;73:S75–S82.
146. Hammond EH, Flinner RL. Clinically relevant breast cancer reporting: using process measures to improve anatomic pathology reporting. *Arch Pathol Lab Med*. 1997;121:1171–1175.
147. Simpson RW, Berman MA, Foulis PR, et al. Cancer biomarkers: the role of structured data reporting. *Arch Pathol Lab Med*. 2015;139:587–593.
148. Roche Diagnostics. VENTANA PD-L1 (SP263) Assay Staining of non-Small Cell Lung Cancer Interpretation Guide; 2016. Available at: www.ventana.com/ventana_portal/OpenOverlayServlet?launchIndex=1&objectId=741-48601015703EN. Accessed February 17, 2019.
149. Dako. PD-L1 IHC 28-8 pharmDx Interpretation Manual—Melanoma; 2016. Available at: www.dako.com/cs/library/usermanuals/public/29120_pd-l1-ihc-28-8-interpretation-manual-melanoma.pdf. Accessed February 17, 2019.
150. Dako. PD-L1 IHC 28-8 pharmDx Interpretation Manual—Non-Squamous Non-Small Cell Lung Cancer; 2015. Available at: www.dako.com/cs/library/usermanuals/public/29111_pd-l1-ihc-28-8-interpretation-manual.pdf. Accessed February 17, 2019.
151. Dako. PD-L1 IHC 22C3 pharmDx Interpretation Manual—Urothelial Carcinoma; 2018. Available at: www.dako.com/cs/library/usermanuals/public/29276_22c3_pharmdx_uc_interpretation_manual_us.pdf. Accessed February 17, 2019.
152. Dako. PD-L1 IHC 28-8 pharmDx Interpretation Manual—Squamous Cell Carcinoma of the Head and Neck; 2017. Available at: www.dako.com/cs/library/usermanuals/public/29191_pd-l1-ihc-28-8-scch-interpretation-manual.pdf. Accessed February 17, 2019.
153. Dako. PD-L1 IHC 22C3 pharmDx Interpretation Manual—Gastric or Gastroesophageal Junction Adenocarcinoma; 2017. Available at: www.dako.com/cs/library/usermanuals/public/29219_pd-l1-ihc-22c3-pharmdx-gastric-interpretation-manual_us.pdf. Accessed February 17, 2019.
154. Dako. PD-L1 IHC 22C3 pharmDx Interpretation Manual—Cervical Cancer; 2018. Available at: www.dako.com/cs/library/usermanuals/public/29257_22c3_pharmdx_cervical_interpretation_manual_us.pdf. Accessed February 17, 2019.
155. Roche Diagnostics. VENTANA PD-L1 (SP142) Assay Interpretation Guide for NSCLC; 2018. Available at: http://productlibrary.ventana.com/ventana_portal/OpenOverlayServlet?launchIndex=1&objectId=741-48601015703EN. Accessed February 17, 2019.
156. Roche Diagnostics. VENTANA PD-L1 (SP142) Assay Interpretation Guide for Urothelial Carcinoma; 2018. Available at: http://productlibrary.ventana.com/ventana_portal/OpenOverlayServlet?launchIndex=1&objectId=741-48601015704EN. Accessed February 17, 2019.
157. Dako. PD-L1 IHC 28-8 pharmDx Interpretation Manual—Urothelial Carcinoma; 2017. Available at: www.dako.com/cs/library/usermanuals/public/29188_pd-l1-ihc-28-8-uc-interpretation-manual.pdf. Accessed February 18, 2019.
158. Giunchi F, Degiovanni A, Daddi N, et al. Fading with time of PD-L1 immunoreactivity in non-small cells lung cancer tissues: a methodological study. *Appl Immunohistochem Mol Morphol*. 2018;26:6.
159. Cheung CC, Banerjee D, Barnes PJ, et al. Canadian Association of Pathologists-Association Canadienne Des Pathologistes National Standards Committee for High Complexity Testing/Immunohistochemistry: guidelines for the preparation, release, and storage of unstained archived diagnostic tissue sections for immunohistochemistry. *Am J Clin Pathol*. 2014;142:629–633.
160. Cheung CC, D'Arrigo C, Diemel M, et al. Evolution of quality assurance for clinical immunohistochemistry in the era of precision medicine: part 4. *Appl Immunohistochem Mol Morphol*. 2017;25:227–230.
161. US Food and Drug Administration. Overview of IVD Regulation; 2018. Available at: www.fda.gov/medicaldevices/deviceregulationandguidance/ivdregulatoryassistance/ucm123682.htm. Accessed January 17, 2018.
162. Cheung CC, Taylor CR, Torlakovic EE. An audit of failed immunohistochemical slides in a clinical laboratory: the role of on-slide controls. *Appl Immunohistochem Mol Morphol*. 2015;25:308–312.
163. Torlakovic EE, Francis G, Garratt J, et al. Standardization of negative controls in diagnostic immunohistochemistry: recommendations from the international ad hoc expert panel. *Appl Immunohistochem Mol Morphol*. 2014;22:241.
164. Dako PD-L1 IHC 22C3 pharmDx; 2018. Available at: www.agilent.com/en-us/products/pharmdx/pd-l1-ihc-22c3-pharmdx-testing#pink3. Accessed August 30, 2018.
165. Terry J, Torlakovic EE, Garratt J, et al. Implementation of a Canadian external quality assurance program for breast cancer biomarkers: an initiative of Canadian Quality Control in immunohistochemistry (ciQc) and Canadian Association of Pathologists (CAP) National Standards Committee/Immunohistochemistry. *Appl Immunohistochem Mol Morphol*. 2009;17:375–382.
166. Vyberg M, Nielsen S. Proficiency testing in immunohistochemistry—experiences from Nordic Immunohistochemical Quality Control (NordiQC). *Virchows Arch*. 2016;468:19–29.
167. Nielsen S. External quality assessment for immunohistochemistry: experiences from NordiQC. *Biotech Histochem*. 2015;90:331–340.
168. UK NEQAS. PD-L1 report of the pilot run; 2017. Available at: www.ukneqasiccish.org/wp/wp-content/uploads/2017/09/PD-L1_prepilot_write_up_070917.pdf. Accessed February 17, 2019.
169. Canadian Association of Pathologists-Association Canadienne Des Pathologistes National Standards Committee, Torlakovic EE, Riddell R, Banerjee D, et al. Canadian Association of Pathologists-Association Canadienne Des Pathologistes National Standards Committee/Immunohistochemistry: best practice recommendations for standardization of immunohistochemistry tests. *Am J Clin Pathol*. 2010;133:354–365.
170. Janes H, Brown MD, Huang Y, et al. An approach to evaluating and comparing biomarkers for patient treatment selection. *Int J Biostat*. 2014;10:99–121.
171. Makretsov N, Gilks CB, Alaghebandan R, et al. Development of an evidence-based approach to external quality assurance for breast cancer hormone receptor immunohistochemistry: comparison of reference values. *Arch Pathol Lab Med*. 2011;135:874–881.
172. International Organization for Standardization/International Electrotechnical Commission (ISO/IEC). Conformity assessment—general requirements for proficiency testing. Available at: www.iso.org/obp/ui/ft/#iso:std:iso-iec:17043:en. Accessed January 28, 2019.
173. Aarsand AK, Sandberg S. How to achieve harmonisation of laboratory testing—the complete picture. *Clin Chim Acta*. 2014;432:8–14.