European Journal of Cancer 106 (2019) 234–243

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejcancer.com

Original Research

Performance of the Food and Drug Administration/EMA-approved programmed cell death ligand-1 assays in urothelial carcinoma with emphasis on therapy stratification for first-line use of atezolizumab and pembrolizumab

Markus Eckstein [a,*,1], Philipp Erben [b,1], Maximilian C. Kriegmair [b,1], Thomas S. Worst [b,1], Cleo-Aron Weiß [c], Ralph M. Wirtz [d,1], Sven Wach [e,1], Robert Stoehr [a,1], Danijel Sikic [e,1], Carol I. Geppert [a], Veronika Weyerer [a,1], Simone Bertz [a], Johannes Breyer [f,1], Wolfgang Otto [f,1], Bastian Keck [e,1], Maximilian Burger [f,1], Helge Taubert [e,1], Wilko Weichert [g,1], Bernd Wullich [e,1], Christian Bolenz [h,1], Arndt Hartmann [a,1], Franziska Erlmeier [g,1]

[a] *Institute of Pathology, University Hospital Erlangen, Friedrich-Alexander University Erlangen-Nürnberg, Erlangen, Germany*
[b] *Department of Urology Mannheim, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany*
[c] *Institute of Pathology, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany*
[d] *STRATIFYER Molecular Pathology GmbH, Cologne, Germany*
[e] *Department of Urology and Pediatric Urology, Friedrich-Alexander University Erlangen-Nürnberg, Erlangen, Germany*
[f] *Department of Urology, University of Regensburg, Regensburg, Germany*
[g] *Institute of Pathology, Technical University Munich, Munich, Germany*
[h] *Department of Urology, University of Ulm, Ulm, Germany*

Received 18 August 2018; received in revised form 4 November 2018; accepted 6 November 2018
Available online 5 December 2018

* *Corresponding author*: Institute of Pathology, University Hospital Erlangen, Friedrich-Alexander University Erlangen-Nürnberg, Krankenhausstraße 8-10.
E-mail addresses: markus.eckstein@uk-erlangen.de (M. Eckstein), Philipp.erben@medma.uni-heidelberg.de (P. Erben), maximilian.kriegmair@umm.de (M.C. Kriegmair), thomas.worst@umm.de (T.S. Worst), cleo-aron.weis@umm.de (C.-A. Weiß), ralph.wirtz@stratifyer.de (R.M. Wirtz), sven.wach@uk-erlangen.de (S. Wach), Robert.stoehr@uk-erlangen.de (R. Stoehr), Danijel.sikic@uk-erlangen.de (D. Sikic), carol.geppert@uk-erlangen.de (C.I. Geppert), veronika.weyerer@uk-erlangen.de (V. Weyerer), simone.bertz@uk-erlangen.de (S. Bertz), johannes.breyer@ukr.de (J. Breyer), wolfgang.otto@ukr.de (W. Otto), bastian.keck@uk-erlangen.de (B. Keck), maximilian.burger@ukr.de (M. Burger), helge.taubert@uk-erlangen.de (H. Taubert), wilko.weichert@tum.de (W. Weichert), bernd.wullich@uk-erlangen.de (B. Wullich), Christian.bolenz@uniklinik-ulm.de (C. Bolenz), Arndt.hartmann@uk-erlangen.de (A. Hartmann), f.erlmeier@tum.de (F. Erlmeier).
[1] On behalf of the BRIDGE-Consortium Germany.

https://doi.org/10.1016/j.ejca.2018.11.007

KEYWORDS
Bladder cancer;
Urothelial carcinoma;
PD-L1 assay;
PD-L1;
PD-1;
Checkpoint inhibition

Abstract ***Background:*** Recently, the Food and Drug Administration (FDA)/European Medicines Agency (EMA) restricted first-line use of atezolizumab and pembrolizumab in patients with metastasised urothelial carcinoma by defining distinct programmed cell death ligand-1 cut-offs. We analysed the diagnostic performance of all FDA/EMA-approved programmed cell death ligand-1 assays with emphasis on new restrictions for first-line treatment with atezolizumab and pembrolizumab.

Patients and methods: Two hundred fifty-one urothelial carcinomas were analysed on tissue microarrays with four cores of each tumour. Stains were performed in certified laboratories on Ventana Benchmark Ultra and Dako Link 48 autostainers. Stains were read on an assay-by-assay basis by two trained pathologists. Overall percentage agreement (OPA) was calculated across the preset cut-offs. Positive percentage agreement (PPA) and negative percentage agreement (NPA) were calculated across different scoring algorithms. Venn diagrams were constructed to illustrate discordance according to the recent FDA/EMA guidelines.

Results: The Dako 28-8, 22c3 and the Ventana SP263 assays showed high interassay correlation (r-range 0.83–0.91). Interassay correlation between the Ventana SP142 and the three other assays was moderate (r-range 0.66–0.75). OPA of 93.3% was achieved between the Dako 28-8, 22c3 and Ventana SP263 assays. OPA including the SP142 was 84.1%. Pooled PPA and NPA of different scoring algorithms was 89.4% and 95.3% for the Dako 28-8, 22c3 and the SP263 assays, respectively. With the SP142 assay, pooled PPA was 59.1%. The SP142 assay identifies fewer eligible patients for first-line treatment with atezolizumab/pembrolizumab.

Conclusion: Dako 28-8, 22c3 and SP263 assays show interchangeable performance. The SP142 assay shows divergent staining results. Interassay variability leads to different detection rates of eligible patients for first-line treatment with atezolizumab and pembrolizumab.

1. Introduction

Exploiting of inhibitory checkpoint proteins such as programmed cell death-1 (PD-1), programmed cell death ligand-1 (PD-L1) or cytotoxic T-lymphocyte–associated protein 4 that suppress antitumour T-cell responses is a common immune-evasive strategy of several solid tumours such as non–small-cell lung cancer (NSCLC), malignant melanoma or urothelial carcinoma (UC) [1–3]. Immunotherapeutic agents targeting PD-1 and PD-L1 are emerging in UC [4–11]. Similar to other solid tumours, higher objective response rates (ORRs) were observed in patients with high expression of PD-L1 [5–7,10,11]. All used drugs are the European Medicines Agency (EMA)/Food and Drug Administration (FDA) approved in conjunction with a complementary PD-L1 assay of which each bases on different antibody clones, immunohistochemistry (IHC) protocols, scoring algorithms, and cut-offs for PD-L1 positivity (Table S2) [6,7]. Recently, first-line use of atezolizumab (IMvigor130) and pembrolizumab (Keynote-361) in platinum-ineligible patients has been restricted to patients with PD-L1–positive tumours due to higher mortality rates of patients with PD-L1–negative tumours receiving checkpoint inhibition instead of chemotherapy [12,13]. Thus, the PD-L1 status assessment in UC might support clinicians to identify patients with improved therapy responsiveness (second line) and protecting patients with PD-L1–negative tumours from harmful effects (first-line use) [12–14].

Most pathology laboratories are not able to buy expensive autostainers from different companies (Ventana/Dako) which are needed to perform the EMA/FDA-approved assays. Therefore, most routine laboratories use to create cheaper laboratory-developed assays which are not validated and approved for the specific indications. This is particularly critical against the background of the current FDA/EMA restriction: both approved assays are each marketed by two different manufacturers (pembrolizumab/Dako 22c3; atezolizumab/Ventana SP142). Because these assays are expensive and not reimbursed by health insurances in several EU countries, this could lead to a widespread withholding of a potentially effective therapy. Furthermore, prior studies demonstrated that PD-L1 assays are not standardised, leading to different staining results potentially affecting appropriate treatment selection [12–16]. Interobserver effects might further affect appropriate therapy selection [17,18].

Taken together, it is necessary to compare the analytic performance of PD-L1 diagnostic assays to allow appropriate interpretation of their use with respect to treatment selection in UC. We conducted this study to investigate the analytic performance of four FDA/EMA-approved assays (Ventana SP142 and SP263, Dako 28-8 and 22c3) in UC with emphasis on

implications for the therapy selection for first-line use of atezolizumab/pembrolizumab.

2. Patients, materials and methods

2.1. Tumour samples and tissue microarray construction

Two hundred fifty-one consecutively collected formalin-fixed, paraffin-embedded UC samples were obtained from two pathologies (treated between 2004 and 2016). All specimens were re-evaluated according to the latest World Health Organization classification (2016) and the latest Union internationale contre le cancer (UICC) staging manual (2017) by three pathologists (A.H./M.E./F.E.). The study was approved by the local ethics committee (Grant:329_16B) and carried out according to the Declaration of Helsinki.

Tissue microarrays (TMAs) were constructed as reported previously [18]. Hematoxylin Eosin (HE) stained slides were scanned (Panoramic P250, 3DHistech) and annotated (CaseViewer v1.0, 3DHistech). Four cores (2× tumour-centre, 2× invasion-front; diameter 1 mm) were taken using an automated tissue microarrayer (TMA-Grandmaster, 3DHistech).

2.2. PD-L1 assays and slide reading

IHC stains were performed on consecutively cut 4-µm sections on a Ventana Benchmark Ultra (SP263/SP142; Ventana, USA) and a Dako Link 48 autostainer (22c3/28-8; Dako, USA) according to the manufacturers protocols in IHC laboratories with accreditation by the German Accreditation Office (DAKKs) according to DIN EN ISO/IEC 17020 (Table S2). Stained TMA sections were scored discretely on an assay-by-assay basis by two trained independent pathologists (M.E., F.E.) according to the current assay recommendations. Per antibody assay, 1004 cores were analysed.

2.3. Statistical analyses

To compare agreement of different assays at different cut-offs, overall percentage agreement (OPA) was calculated pairwise. To compare concordant and discordant classification among different assays at prespecified cut-offs, positive percentage agreement (PPA) and negative percentage agreement (NPA) were calculated. The prespecified antibody assay was set as reference for calculation of PPA/NPA. Mean values of immune cell (IC), tumour cell (TC) and combined positive score (CPS) staining were tested by a non-parametric Wilcoxon test (Fig. 1C). Venn diagrams were constructed to illustrate the discordance/concordance of different assays and cut-off systems. Kappa values were calculated to illustrate interobserver.

Statistical analyses and graphical visualisations were performed by JMP SAS 13.2 (SAS, Cary, USA) and GraphPad Prism 7.03 (GraphPad; USA). All tests were two sided, and p-values of <0.05 were considered to be significant.

3. Results

3.1. Clinicopathological data

Clinicopathological data are shown in Table S1. All patients had muscle-invasive UC of the bladder (pT2a-pT4b). Median age was 71 years (minimum: 37; maximum: 91). None of the patients received checkpoint inhibition.

3.2. PD-L1 TC and IC staining

The Dako 22c3, 28-8 and the Ventana SP263 assays showed similar amounts of positively detected TC and IC, whereas the SP142 assay detected especially fewer TC but also fewer IC (Fig. 1A–C). Staining intensity of the 28-8 and 22c3 assays was weaker compared with SP263 and SP142. The SP142 assay showed a typical ant- or dot-like IC and TC staining pattern (Fig. 1A/B).

3.3. Interassay correlation

Intercore and inter-region (tumour centre/invasive margin) agreements are shown in Table S3. Non-parametric Spearman-rank correlations between assays are shown in Fig. S1. Correlation between the Dako 28-8, 22c3 and the Ventana SP263 assays ranges from 0.86 to 0.91 for TC and from 0.83 to 0.90 for IC. Interassay correlation including the SP142 assay ranges from 0.66 to 0.72 for TC and from 0.69 to 0.75 for IC, revealing moderate associations between the other assays and the SP142 assay (Fig. S1).

3.4. OPA across various objective and prespecified scoring algorithm cut-offs

Pooled OPA was 93.3% (range: 88.0–98.0%) between the Dako 22c3, 28-8 and the Ventana SP263 assays at multiple cut-offs, indicating high interassay agreement and interchangeability (Table S4; Fig. 1D). Inclusion of the SP142 assay (SP142 vs. 22c3/28-8/SP263) revealed pooled OPA of 84.1% (range: 75.6–98.4%), indicating comparability, but not interchangeability (Table S4; Fig. 1D).

3.5. PPA and NPA across prespecified scoring algorithm cut-offs.

Pooled OPA across prespecified algorithm cut-offs was 92.7% (range: 88.4%–96.4%) between the Dako 22c3, 28-8 and the SP263 assays (Fig. 1E; Table S4). Compared with the SP142 assay, pooled OPA across prespecified algorithm cut-offs was 82.0% (range: 77.3–88.0%). Pooled PPA and NPA of different scoring

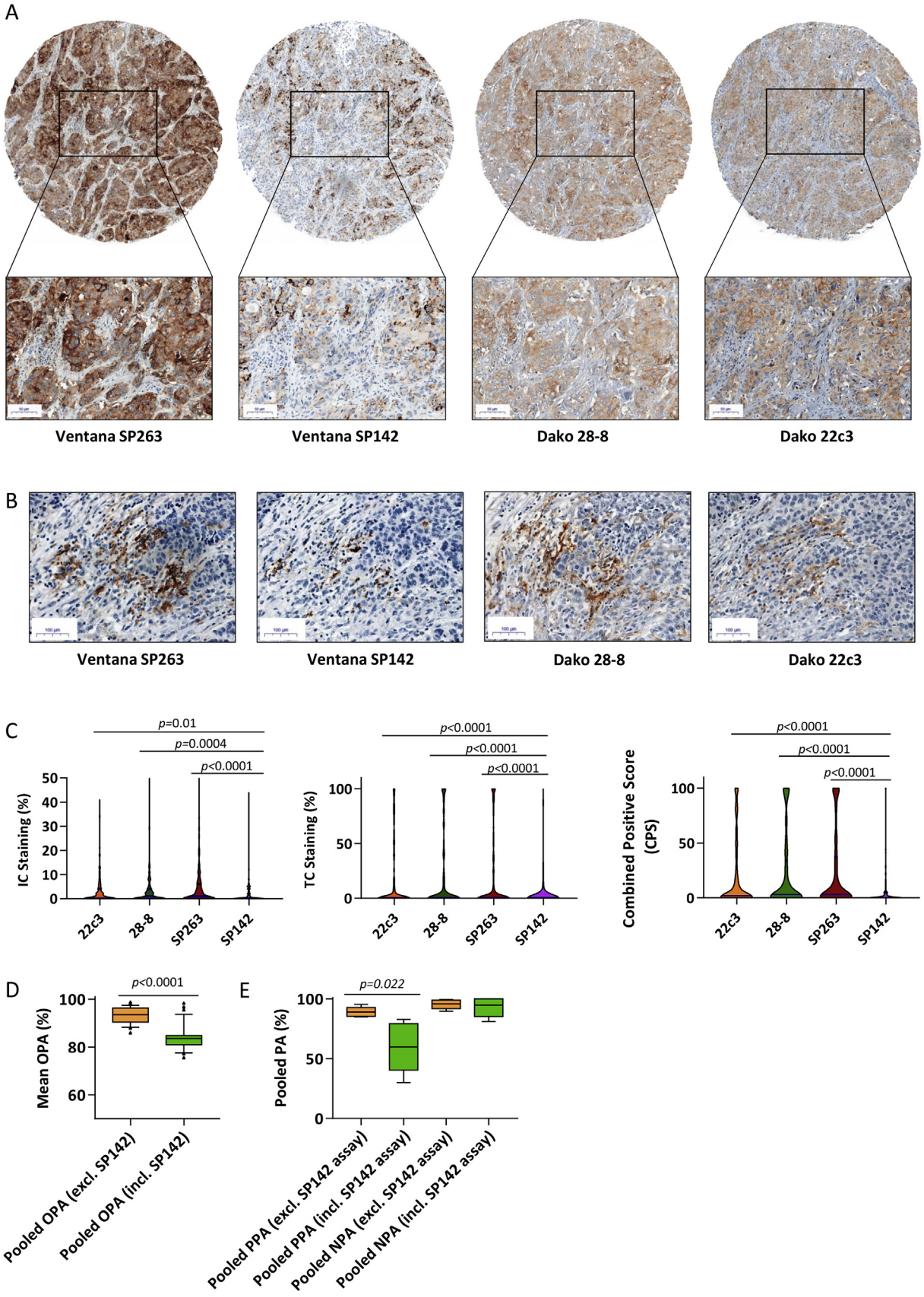

Fig. 1. (A) Representative images of tumour cell (TC) membrane staining of all four applied assays. (B) Representative images of immune cell (IC) staining of all four applied assays. (C) Overall positive staining of ICs and TCs across the four assays. (D) Pooled overall percentage agreement (OPA) between the Dako and the SP263 assay ('excl. SP142 assay') and between the Dako/SP263 assays and SP142 assay ('incl. SP142 assay'). (E) Pooled positive percentage agreement (PPA) and pooled negative percentage agreement (NPA) between the Dako and the SP263 ('excl. SP142 assay') assays and between the Dako/SP263 assays and the SP142 assay ('incl. SP142 assay').

Table 1
Positive percentage agreement (PPA) and negative percentage agreement (NPA) in 251 muscle-invasive bladder cancer specimens across different PD-L1 assay algorithms.

Reference assay	Comparator assay at matched expression cut-off							
	Dako 28-8		Dako 22c3		Ventana SP142		Ventana SP263	
	PPA	NPA	PPA	NPA	PPA	NPA	PPA	NPA
DAKO 28-8 </≥ 5% TC			88.1%	96.4%	44.0%	99.4%	85.0%	95.2%
DAKO 22c3 </≥ 10 CPS	95.5%	92.6%			46.6%	100%	92.0%	89.6%
Ventana SP142 </≥ 5% IC	73.2%	86.7%	78.0%	90.0%			82.9%	81.0%
Ventana SP263 </≥ 25% TC- or IC	90.0%	98.8%	85.7%	99.4%	30.0%	100%		

TC, tumour cells; IC, immune cells; CPS, combined positive score.
PPA ≥ 85% indicates interchangeability of diagnostic assays.

algorithms was 89.1% and 95.8% for the Dako 22c3, 28-8 and the SP263 assays, respectively. Compared with the SP142 assay (SP142 vs. Dako 22c3, 28-8 and Ventana SP263 assays), pooled PPA was 59.9% and pooled NPA 94.7% (Table 1, Fig. 1E). The positivity rate of all cases at specific cut-offs used in clinical trials dependent on different antibody assays are depicted in Table 2.

3.6. Single value agreement across different assays

Single scoring values of IC, TC and CPS showed a huge overlap for most analysed samples across the Dako 22c3, 28-8 and Ventana SP263 assays (Fig. 2). The Ventana SP142 assay detected especially fewer TCs (Fig. 2).

3.7. Influence of formalin-fixed paraffin embedded (FFPE) tissue age on PD-L1 staining

Differences in PD-L1 staining seemed to be rather influenced by utilised antibody assays than by tissue age (Fig. 2D).

3.8. Agreement across different clinical important cut-offs for first-line therapy selection with atezolizumab and pembrolizumab

Regarding the current FDA/EMA restrictions for first-line use of atezolizumab (IC ≥ 5% per tumour area), the Dako 22c3 (1.29-fold), Dako 28-8 (1.41-fold) and Ventana SP263 (1.8-fold) assays identified significantly more cases as PD-L1 positive than the reference assay (Ventana SP142; Fig. 3A). Vice versa, regarding the current FDA/EMA regulations for the first-line usage of pembrolizumab (CPS≥10), the SP142 assay identified significantly less tumours as positive (0.47-fold) than the reference assay (Dako 22c3), while the two other assays showed similar but slightly higher positive classification rates than the reference assay (DAKO 28-8 1.09-fold, Ventana SP263 1.11-fold; Fig. 3A).

The Dako 22c3, 28-8 and the Ventana SP263 assays showed pooled OPA of 90.4% for the 5%-IC cut-off and 92.0% for the CPS≥10 cut-off (Table S4). Regarding both cut-offs, all FDA/EMA-approved assays exclusively identified several cases as positive, while the other assays did not. Despite wide interchangeability in terms of OPA, PPA and NPA (22c3, 28-8, SP263), this underlines a remaining and clinically relevant interassay variability (Fig. 3B). Overall, the SP142 assay showed the lowest positive classification rate with 16.3% for both cut-off systems (Fig. 3A/B). Compared with atezolizumab trials, the SP142 5%-IC positivity rate is lower but congruent with prior independent reports [19].

To analyse the concordance or discordance between the CPS10- and IC5% cut-off system, we constructed a Venn diagram. Cases were defined as positive if the PD-L1 expression value was equal or above the required cut-off in at least one assay to analyse exclusively the inter-algorithm variability, independent of interassay variabilities (summarised positivity classification rate; Table

Table 2
Positivity rates of consensus scoring between both observers of cases at different cut-off levels used in clinical trials.

Positivity rate (n, %)	TC/IC ≥ 25% [6]	IC ≥ 5% [7]	CPS≥10 [4]	TC ≥ 5% [24]
Summary	73 (29.1%)	88 (35.1%)	110 (43.8%)	95 (37.8%)
Ventana SP263	70 (27.9%)	74 (29.5%)	98 (39.0%)	79 (31.5%)
Ventana SP142	21 (8.4%)	41 (16.3%)	41 (16.3%)	38 (15.1%)
Dako 22c3	61 (24.3%)	53 (21.1%)	88 (35.1%)	80 (31.9%)
Dako 28-8	65 (25.9%)	58 (23.1%)	96 (38.2%)	84 (33.5%)

TC, tumour cells; IC, immune cells; CPS, combined positive score.
'Summary' includes all cases classified as positive by at least by one diagnostic assay.

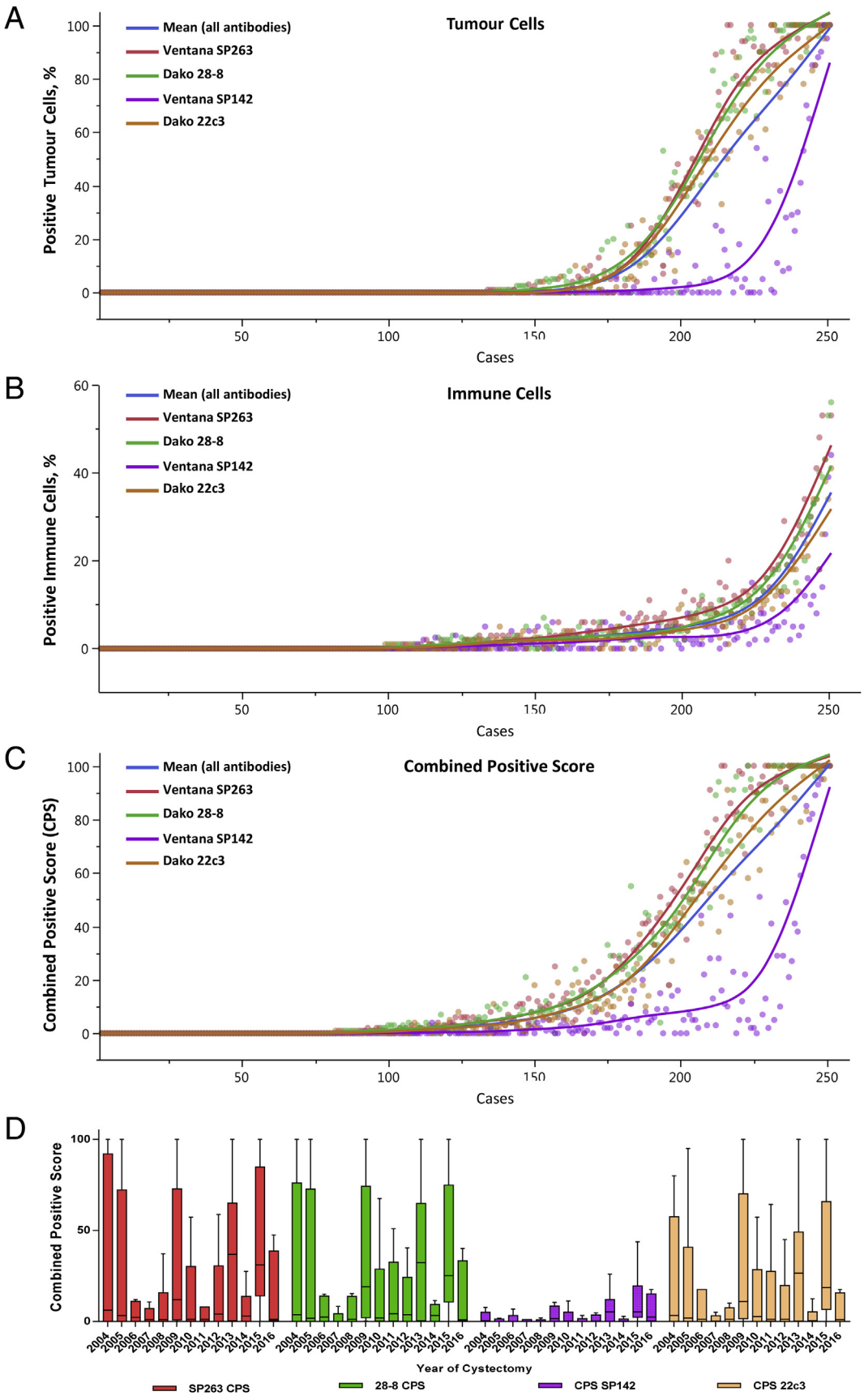

Fig. 2. (A) Tumour cell proportion scores from all read samples across the four different assays. (B) Immune cell proportion scores from all read samples across the four different assays. (C) Combined positive score (CPS) from all read samples across the four different assays. (D) Correlation of tissue age (age of FFPE blocks) and PD-L1 staining (CPS). PD-L1, programmed cell death ligand-1; FFPE, formalin-fixed paraffin embedded.

A

PD-L1 Diagnostic Assays (Atezolizumab)	IC ≥5%	IC <5%	Fold-Change (positive cases)
Reference Assay: Ventana SP142	41 (16.3%)	210 (83.7%)	Reference
Ventana SP263	74 (29.5%)	177 (70.5%)	**1.80**
DAKO 22c3	53 (21.1%)	198 (78.9%)	**1.29**
DAKO 28-8	58 (23.1%)	193 (76.9%)	**1.41**
PD-L1 Diagnostic Assays (Pembrolizumab)	**CPS ≥10**	**CPS <10**	**Fold-Change (positive cases)**
Reference Assay: DAKO 22c3	88 (35.0%)	163 (65.0%)	Reference
DAKO 28-8	96 (38.2%)	155 (61.7%)	**1.09**
Ventana SP263	98 (39.0%)	153 (61.0%)	**1.11**
Ventana SP142	41 (16.3%)	210 (83.7%)	**0.47**

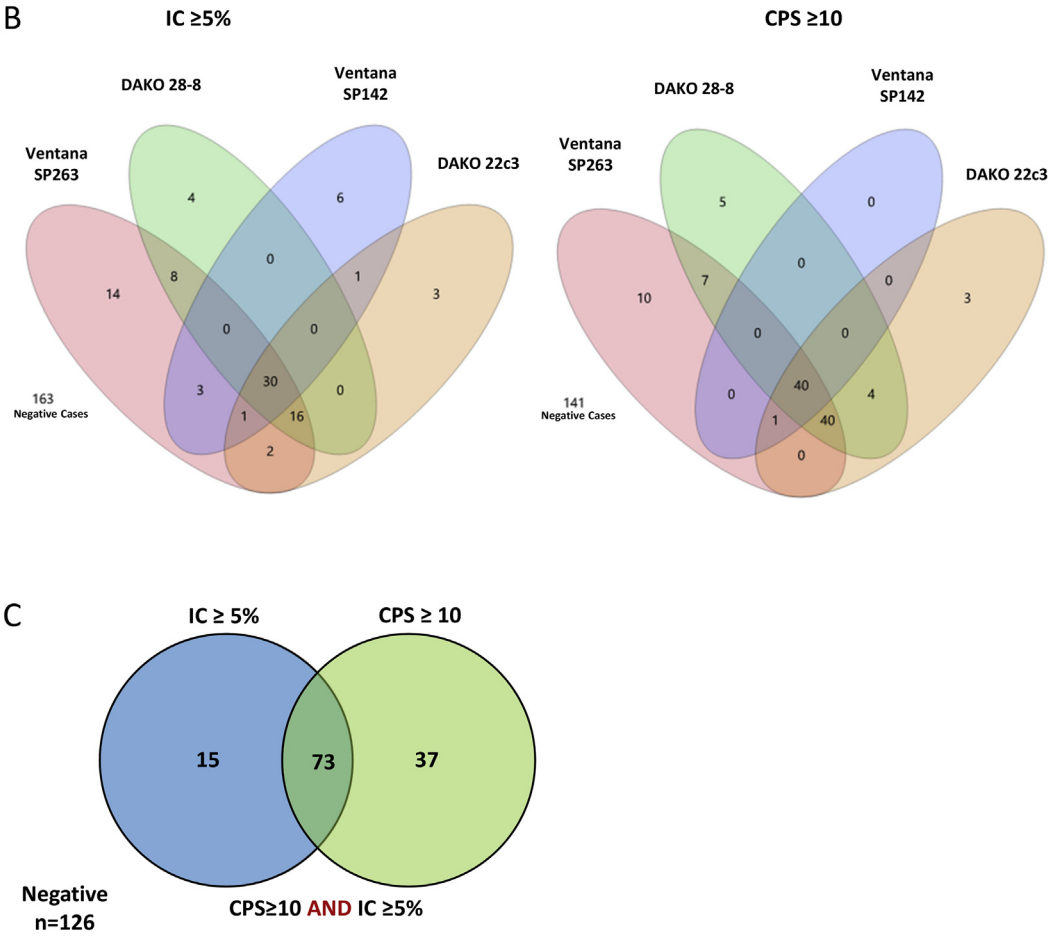

Fig. 3. (A) Amounts of patients identified as eligible for first-line treatment with atezolizumab (≥5% immune cells [ICs]) and pembrolizumab (≥10 combined positive score [CPS]) according to the recently published restrictions of the FDA and EMA for the first-line use of atezolizumab and pembrolizumab. (B) Venn diagrams illustrating results across different assays according to the current FDA/EMA restrictions. (C) Venn diagram illustrating discordant classifications of the two clinical relevant scoring algorithms: CPS10 vs. 5%-IC. Cases were classified as positive if the PD-L1 expression value exceeded the respective cut-off value in at least one diagnostic PD-L1 assay to exclude the influence of interassay variabilities. FDA, Food and Drug Administration; PD-L1, programmed cell death ligand-1.

2). One hundred twenty-five cases (49.8%) were identified as positive. Seventy-three (29.1%) cases were concordantly positive, while 15 cases (6.0%) were exclusively CPS10-positive and 37 (14.7%) exclusively above the 5%-IC cut-off (Fig. 3C).

Interobserver agreement of both observers was substantial to almost perfect (κ-range: 0.73–0.90; Table 3).

4. Discussion

In UC, most clinical trials observed higher ORR in patients with high PD-L1 expression but diverging results concerning progression-free and overall survival [5–7,10,11]. Reasons for this 'consistent' inconsistency are not yet clarified, but one major aspect might be the

Table 3
Kappa values of the two trained observers.

Assay	TC/IC ≥ 25% [6]	IC ≥ 5% [7]	CPS≥10 [4]	TC ≥ 5% [24]
Ventana SP263	0.90 [0.84–0.96]	0.76 [0.67–0.85]	0.87 [0.81–0.94]	0.84 [0.77–0.91]
Ventana SP142	0.85 [0.72–0.96]	0.76 [0.64–0.87]	0.86 [0.78–0.95]	0.83 [0.73–0.92]
Dako 22c3	0.88 [0.81–0.94]	0.73 [0.63–0.84]	0.85 [0.79–0.92]	0.83 [0.76–0.91]
Dako 28–8	0.89 [0.83–0.95]	0.75 [0.65–0.84]	0.84 [0.77–0.91]	0.81 [0.74–0.89]

TC, tumour cells; IC, immune cells; CPS, combined positive score.
95% confidence intervals are depicted in dashes.

inter–PD-L1 assay/algorithm variability [17,20]. Recently, the FDA and EMA restricted the first-line use of atezolizumab and pembrolizumab in patients with low PD-L1 expression due to increased mortality rates in the PD-L1–low/negative subpopulation. Therefore, IHC PD-L1 assessment is now obligate for platinum ineligible patients before first-line use of atezolizumab or pembrolizumab [12,13]. Because numerous phase III/IV trials are still ongoing, it still remains unclear whether PD-L1 assessment will become obligate in further indications. This development is particularly important against the background that every single agent is approved in conjunction with a specific premade assay which has to be performed on a specific staining platform (Dako/Ventana). Because autostainer platforms and the premade assays are expensive, it is likely that most laboratories substitute approved assays by a single (laboratory developed) assay which harbours a potential risk of misclassifying compared with premade, validated assays due to known effects of interlaboratory and interassay variability. Potential interchangeability of different assays would enable laboratories to use only one, standardised assay for general PD-L1 assessment.

In this study, TC and IC staining of Dako 22C3, 28-8 and Ventana SP263 demonstrated interchangeable analytic performance at multiple matched expression cut-offs, whereas the Ventana SP142 assay consistently detected fewer positive TCs/ICs which has been previously demonstrated for NSCLC [15,17,20–22]. The Dako 22c3, 28-8 and Ventana SP263 assays have closely aligned dynamic ranges and showed a pooled OPA of 92.8% across different cut-offs for IC and TC, which is a characteristic for interchangeable IHC histochemical assays [23]. Inclusion of the SP142 assay revealed a lower pooled OPA of 84.1% indicating comparability but not interchangeability [23].

Regarding specific scoring algorithms used within clinical trials and for first-line therapy stratification, the Dako 22c3, 28-8 and Ventana SP263 assays showed interchangeable performance (pooled PPA 89.4%; Table 1). Interestingly, the performance was even better for algorithms with high cut-off values such as the TC/IC 25% algorithm (PPA: 87.8% and durvalumab; [6]) and the CPS algorithm (PPA: 95.8% and pembrolizumab; [4]), indicating that higher cut-off levels are easier to resolute and are likely more robust against interobserver effects which has been demonstrated previously [17,18,20,21]. Compared with the other assays, the performance of the SP142 assay for the 5%-IC was comparable but not interchangeable (pooled PPA: 78.0%; Table 1). Regarding algorithms including TC staining (5% TC, CPS10 and TC/IC 25%), the performance of the SP142 assay was poor with a pooled PPA of 40.2%, indicating that this assay should not be used for these algorithms. Taken together, these data indicate that the intrinsic assay variability between the Dako 22c3, 28-8 and the SP263 assays is low, and therefore, those assays may be used interchangeable for different scoring algorithms, which is congruent with prior data on UC and NSCLC [15,16,20]. Staining characteristics of the SP142 assay are comparable for IC scoring but not interchangeable [23]. Moreover, the SP142 shows poor agreement for algorithms including TC compared with the validated reference assays indicating that this assay should not be used for these algorithms without further validation.

The remaining interassay variability and, especially, the discordances in comparison to the SP142 assay might harbour clinical importance [12–14]. In our study cohort of 251 patients with muscle-invasive bladder cancer (MIBC), the SP142 assay would classify 0.55- to 0.77-fold less patients as eligible for first-line treatment with atezolizumab (5%-IC cut-off; Fig. 3A [7]). Vice versa, the Dako 28-8, 22c3 and Ventana SP263 assays (CPS≥10) would detect more than twice as many eligible patients for first-line treatment with pembrolizumab than the SP142 assay (2.14- to 2.39-fold; Fig. 3A) [4]. Despite theoretical interchangeability, the Dako 28-8 and the Ventana SP263 assays classify more patients as eligible for first-line treatment with pembrolizumab than the 22c3 assay (Fig. 3A). Moreover, relevant amounts of samples are exclusively classified as positive by a specific assay (Fig. 3B). This is particularly important because neither the FDA nor the EMA prescribed a specific assay to assess treatment eligibility. This could lead to critical scenarios, e.g. the same patient could receive first-line treatment if tested with assay A but not if tested with assay B (Fig. 3A/B). Although the SP142 assay shows diverging staining characteristics, it is important to note that this does not necessarily mean a worse patient selection for first-line checkpoint inhibition. The SP142 assay seems to identify fewer but preferentially highly PD-L1–positive tumours with potentially increased responsiveness, whereas the other

assays might detect additional therapy responders with lower PD-L1 expression. However, at the moment, there is no evidence that one scoring algorithm or antibody assay is superior in identifying therapy responders.

Assuming that the approved checkpoint inhibitors have similar efficacy, there is another important point to address: Inter-algorithm variability. The concordant positive classification rate of the CPS10 and 5%-IC cut-off system amounted just 58.4% (n = 73; eligible for both drugs), while 15 (12%) patients would be exclusively eligible for pembrolizumab and 37 (29.6%) exclusively for atezolizumab (Fig. 3C). This could lead to withholding of a potentially effective therapy, for example, if only one of the drugs is locally available. A further influencing factor on therapy selection is interobserver variability. Interobserver agreement between the two trained observers in our study was substantial to almost perfect and slightly better for scoring algorithms with high cut-off values. This underlines the importance of systematic training for pathologists. However, larger interobserver studies are needed to investigate real-life interobserver agreement.

Our study demonstrates large analytic concordance between three assays (22c3/28-8/SP263), while the SP142 assay shows comparable—but not interchangeable—results for IC, but not for TC scoring. These analytical and inter-algorithm variabilities led to clinically relevant discordant classifications in a hypothetic model of first-line therapy stratification for the use of pembrolizumab/atezolizumab which could lead to a potential withholding of an effective therapy option.

5. Limitations

A limitation of this study is the lack of outcome data because these patients were not treated with PD-1− or PD-L1−targeted agents. Therefore, we can only evaluate the present data in the context of interassay comparison. However, the distribution of PD-L1 expression in the study population reflects reported results from randomised clinical trials and therefore, also reflects the real-life setting of pretreatment PD-L1 assessment for first-line therapy stratification with atezolizumab/pembrolizumab. Further limitations are the use of TMAs and the analysis of old tissue blocks which could lead to issues due to tumour heterogeneity and age-dependent loss of antigenicity.

Conflicts of interest statement

All authors declare that there is no conflict of interest.

Funding

No third-party funding was received to finance this study.

Research funding

The study was funded by the common scientific pool of every participating institution. There was no writing assistance.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejca.2018.11.007.

References

[1] Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. Nat Rev Cancer 2012;12:252–64.
[2] Postow MA, Callahan MK, Wolchok JD. Immune checkpoint blockade in cancer therapy. J Clin Oncol 2015;33:1974–82.
[3] Bellmunt J, Mullane SA, Werner L, Fay AP, Callea M, Leow JJ, et al. Association of PD-L1 expression on tumor-infiltrating mononuclear cells and overall survival in patients with urothelial carcinoma. Ann Oncol 2015;26:812–7.
[4] Bellmunt J, de Wit R, Vaughn DJ, Fradet Y, Lee JL, Fong L, et al. Pembrolizumab as second-line therapy for advanced urothelial carcinoma. N Engl J Med 2017;376:1015–26.
[5] Balar AV, Galsky MD, Rosenberg JE, Powles T, Petrylak DP, Bellmunt J, et al. Atezolizumab as first-line treatment in cisplatin-ineligible patients with locally advanced and metastatic urothelial carcinoma: a single-arm, multicentre, phase 2 trial. Lancet 2017;389:67–76.
[6] Massard C, Gordon MS, Sharma S, Rafii S, Wainberg ZA, Luke J, 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;34:3119–25.
[7] Rosenberg JE, Hoffman-Censits J, Powles T, van der Heijden MS, Balar AV, Necchi A, 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;387:1909–20.
[8] Balar AV, Castellano D, O'Donnell PH, Grivas P, Vuky J, Powles T, 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;18:1483–92.
[9] Powles T, Duran I, van der Heijden MS, Loriot Y, Vogelzang NJ, De Giorgi U, et al. Atezolizumab versus chemotherapy in patients with platinum-treated locally advanced or metastatic urothelial carcinoma (IMvigor211): a multicentre, open-label, phase 3 randomised controlled trial. Lancet 2018 Feb 24;391(10122):748–57. https://doi.org/10.1016/S0140-6736(17)33297-X. Epub 2017 Dec 18.
[10] Powles T, Eder JP, Fine GD, Braiteh FS, Loriot Y, Cruz C, et al. MPDL3280A (anti-PD-L1) treatment leads to clinical activity in metastatic bladder cancer. Nature 2014;515:558–62.
[11] Powles T, O'Donnell PH, Massard C, Arkenau HT, Friedlander TW, Hoimes CJ, 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.
[12] (FDA) USFaDA, Keytruda (pembrolizumab) or Tecentriq (atezolizumab): FDA Alerts Health Care Professionals and Investigators. FDA statement – decreased survival in some patients in clinical trials associated with monotherapy. 2018.

[13] EMA. EMA restricts use of Keytruda and Tecentriq in bladder cancer data show lower survival in some patients with low levels of cancer protein PD-L1. 2018.
[14] Gourd E. EMA restricts use of anti-PD-1 drugs for bladder cancer. Lancet Oncol 2018 Jul;19(7):e341. https://doi.org/10.1016/S1470-2045(18)30433-9. Epub 2018 Jun 7.
[15] Gaule P, Smithy JW, Toki M, Rehman J, Patell-Socha F, Cougot D, et al. A quantitative comparison of antibodies to programmed cell death 1 ligand 1. JAMA Oncol 2016 Aug 18. https://doi.org/10.1001/jamaoncol.2016.3015 [Epub ahead of print].
[16] Tretiakova M, Fulton R, Kocherginsky M, Long T, Ussakli C, Antic T, et al. Concordance study of PD-L1 expression in primary and metastatic bladder carcinomas: comparison of four commonly used antibodies and RNA expression. Mod Pathol 2018;31:623–32.
[17] Scheel AH, Dietel M, Heukamp LC, Johrens K, Kirchner T, Reu S, et al. Harmonized PD-L1 immunohistochemistry for pulmonary squamous-cell and adenocarcinomas. Mod Pathol 2016;29:1165–72.
[18] Eckstein M, Wirtz RM, Pfannstil C, Wach S, Stoehr R, Breyer J, et al. A multicenter round robin test of PD-L1 expression assessment in urothelial bladder cancer by immunohistochemistry and RT-qPCR with emphasis on prognosis prediction after radical cystectomy. Oncotarget 2018;9:15001–14.
[19] de Jong JJ, Stoop H, Nieboer D, Boormans JL, van Leenders G. Concordance of PD-L1 expression in matched urothelial bladder cancer specimens. Histopathology 2018 Dec;73(6):983–9. https://doi.org/10.1111/his.13710. Epub 2018 Sep 25.
[20] Ratcliffe MJ, Sharpe A, Midha A, Barker C, Scott M, Scorer P, et al. Agreement between programmed cell death ligand-1 diagnostic assays across multiple protein expression cutoffs in non-small cell lung cancer. Clin Canc Res 2017;23:3585–91.
[21] Rimm DL, Han G, Taube JM, Yi ES, Bridge JA, Flieder DB, 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;3:1051–8.
[22] Hirsch FR, McElhinny A, Stanforth D, Ranger-Moore J, Jansson M, Kulangara K, et al. PD-L1 immunohistochemistry assays for lung cancer: results from phase 1 of the blueprint PD-L1 IHC assay comparison project. J Thorac Oncol 2017;12:208–22.
[23] Fitzgibbons PL, Bradley LA, Fatheree LA, Alsabeh R, Fulton RS, Goldsmith JD, 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;138:1432–43.
[24] Sharma P, Callahan MK, Bono P, Kim J, Spiliopoulou P, Calvo E, et al. Nivolumab monotherapy in recurrent metastatic urothelial carcinoma (CheckMate 032): a multicentre, open-label, two-stage, multi-arm, phase 1/2 trial. Lancet Oncol 2016;17:1590–8.