Baseline Serum Autoantibody Signatures Predict Recurrence and Toxicity in Melanoma Patients Receiving Adjuvant Immune Checkpoint Blockade

Paul Johannet[1], Wenke Liu[2], David Fenyo[2], Megan Wind-Rotolo[3], Michelle Krogsgaard[4], Janice M. Mehnert[1], Jeffrey S. Weber[1], Judy Zhong[5], and Iman Osman[1,6]

ABSTRACT

Purpose: Adjuvant immunotherapy produces durable benefit for patients with resected melanoma, but many develop recurrence and/or immune-related adverse events (irAE). We investigated whether baseline serum autoantibody (autoAb) signatures predicted recurrence and severe toxicity in patients treated with adjuvant nivolumab, ipilimumab, or ipilimumab plus nivolumab.

Experimental Design: This study included 950 patients: 565 from CheckMate 238 (408 ipilimumab versus 157 nivolumab) and 385 from CheckMate 915 (190 nivolumab versus 195 ipilimumab plus nivolumab). Serum autoAbs were profiled using the HuProt Human Proteome Microarray v4.0 (CDI Laboratories, Mayaguez, PR). Analysis of baseline differentially expressed autoAbs was followed by recurrence and severe toxicity signature building for each regimen, testing of the signatures, and additional independent validation for nivolumab using patients from CheckMate 915.

Results: In the nivolumab independent validation cohort, high recurrence score predicted significantly worse recurrence-free survival [RFS; adjusted HR (aHR), 3.60; 95% confidence interval (CI), 1.98–6.55], and outperformed a model composed of clinical variables including PD-L1 expression ($P < 0.001$). Severe toxicity score was a significant predictor of severe irAEs (aHR, 13.53; 95% CI, 2.59–86.65). In the ipilimumab test cohort, high recurrence score was associated with significantly worse RFS (aHR, 3.21; 95% CI, 1.38–7.45) and severe toxicity score significantly predicted severe irAEs (aHR, 11.04; 95% CI, 3.84–37.25). In the ipilimumab plus nivolumab test cohort, high autoAb recurrence score was associated with significantly worse RFS (aHR, 6.45; 95% CI, 1.48–28.02), and high severe toxicity score was significantly associated with severe irAEs (aHR, 23.44; 95% CI, 4.10–212.50).

Conclusions: Baseline serum autoAb signatures predicted recurrence and severe toxicity in patients treated with adjuvant immunotherapy. Prospective testing of the signatures that include datasets with longer follow-up and rare but more severe toxicities will help determine their generalizability and potential clinical utility.

See related commentary by Hassel and Luke, p. 3914

Introduction

Adjuvant immune checkpoint blockade (ICB) produces clinical benefit for patients with resected melanoma, but many suffer recurrent disease (1, 2). In some cases, treatment-related toxicity can be severe enough to necessitate the interruption or permanent discontinuation of immunotherapy and may require the use of systemic immunosup-pression. These immune-related adverse events (irAE) can also lead to lifelong need for medications or, in rare cases, death (3, 4). Thus, there is an unmet clinical need to identify biomarkers of immunotherapy response and toxicity. Ideally, a single assay that simultaneously risk-stratifies patients according to their likelihood of suffering recurrence and developing irAEs would help optimize patient selection for treatment.

We hypothesized that some patients possess a subclinical predisposition for ICB toxicity, which is characterized by the presence of autoantibodies (autoAb) before treatment, and which does not manifest spontaneously, but could be unmasked after ICB. Several studies, including one from our own group, suggest an association between baseline autoAbs and overall or site-specific toxicities, but did not test the association between pretreatment autoAbs and ICB efficacy (5–7). The presence of pretreatment, melanoma-associated antibodies has been linked to better immunotherapy treatment outcomes, but this was in a small cohort of patients (8). In another study, baseline levels of rheumatoid factor correlated with progression-free survival and the development of irAEs after PD-1 blockade in patients with non–small cell lung cancer, but the study did not include patients with melanoma (9).

In this study, we leveraged prospectively collected data and serum samples from two phase III clinical trials to identify, test, and independently validate distinct autoAb signatures that predicted recurrence and severe toxicity for patients treated with adjuvant nivolumab. We report similar observations in test cohorts of patients treated with adjuvant ipilimumab monotherapy or ipilimumab plus nivolumab.

[1]Department of Medicine, NYU Grossman School of Medicine, New York, New York. [2]Department of Biochemistry and Molecular Pharmacology, NYU Grossman School of Medicine, New York, New York. [3]Bristol-Myers Squibb, Princeton, New Jersey. [4]Department of Pathology NYU, Grossman School of Medicine, New York, New York. [5]Department of Population Health, NYU Grossman School of Medicine, New York, New York. [6]Ronald O. Perelman Department of Dermatology, NYU Grossman School of Medicine, New York, New York.

Corresponding Authors: Iman Osman, Department of Dermatology, NYU Grossman School of Medicine, 550 First Avenue, Smilow 403, New York, NY 10016. Phone: 212-263-9075; E-mail: Iman.Osman@nyulangone.org; and Judy Zhong, Department of Population Health, NYU Grossman School of Medicine, 180 Madison Avenue, Room 452, New York, NY 10016. Phone: 646-501-3646; E-mail: Judy.Zhong@nyulangone.org

Translational Relevance

In this study, we analyzed pretreatment serum from 950 patients with resected melanoma from two phase III randomized controlled trials of adjuvant immunotherapy: CheckMate 238 (ipilimumab vs. nivolumab) and CheckMate 915 (nivolumab vs. ipilimumab plus nivolumab). We identified autoantibody (autoAb) signatures for nivolumab, ipilimumab, and ipilimumab plus nivolumab that could be used to predict disease recurrence and severe immune-related adverse events. The composite panel of autoAb signatures can allow for the simultaneous risk stratification of patients according to their likelihood of recurring and suffering severe toxicity.

Materials and Methods

Patient population

This study included 950 patients from two phase III randomized controlled trials. There were 565 patients from CheckMate 238 (NCT02388906), an investigation of adjuvant nivolumab (3 mg/kg every 2 weeks) versus ipilimumab (10 mg/kg every 3 weeks for 4 doses and then every 12 weeks), both for 1 year, in patients with high risk, completely resected, American Joint Committee on Cancer (AJCC) 7th Edition stage IIIB-C or IV melanoma (1). There were 385 patients from CheckMate 915 (NCT03068455), which evaluated adjuvant nivolumab (240 mg every 2 weeks) plus ipilimumab (1 mg/kg every 6 weeks) versus nivolumab (480 mg every 4 weeks), both for 1 year, in patients who underwent complete resection of AJCC 8th Edition stage IIIB-D or IV melanoma (10).

Primary study outcomes

The outcomes of interest were: (i) recurrence versus no recurrence, and (ii) severe (grade 3–5) irAEs versus no or mild (grade 1–2) irAEs. Complete information regarding the assessment of recurrence and toxicity during each trial was previously reported (1, 10). Treatment-related events with a potential immunologic etiology were identified using a list of prespecified terms from the *Medical Dictionary for Regulatory Activities*. Events were categorized on the basis of the organ system of origin. The severity of irAEs was graded using the NCI Common Terminology Criteria for Adverse Events, version 4.0.

Serum autoAb profiling

Peripheral blood samples were prospectively collected within 72 hours before administration of the first dose of study medication and processed as detailed in the trial protocols (1, 10). Serum autoAbs were profiled using the HuProt Human Proteome Microarray v4.0 (CDI Laboratories, Mayaguez, PR) as previously described (7, 11). We reasoned that a high-density platform is required for discovery and that a more restricted array with limited numbers of proteins might close the discovery space. We elected to use the HuProt Microarray because it is an extensive platform that contains over 21,000 unique, individually purified full-length human proteins and protein isoforms in duplicate, covering more than 81% of the proteome. In addition, the HuProt Microarray includes 2,000 proteins that are complete technical replicates of another protein on the array. The net signal was defined as the background-subtracted median intensity of each antigen spot. The net signal was log2 transformed and normalized to the median of the total signal intensities on the array for each subject. This served as an internal control to account for the variety in overall baseline autoAb

levels between different subjects. We then standardized each autoAb intensity by its mean and SD across all subjects.

Recurrence and severe toxicity signature development

CheckMate 238 patients who received nivolumab were randomly assigned to training and test sets in a ratio of 75% to 25%. The nivolumab patients from CheckMate 915 served as an independent validation dataset. Ipilimumab patients from CheckMate 238 and ipilimumab plus nivolumab patients from CheckMate 915 were randomly assigned into training and test sets in a ratio of 75% to 25%. Patients' demographic and clinic characteristics were compared between the training, test, and independent validation datasets. The signatures to predict recurrence and severe toxicity were derived from the training sets for each regimen. For each outcome in each regimen, we first identified a subset of autoAbs that: (i) had higher intensities in the no-recurrence group than in the recurrence group or in the severe-toxicity compared with the no-severe-toxicity group, and (ii) had a univariate P value less than 0.05 as determined by univariable two-sample t test comparing the recurrence versus no recurrence groups or severe toxicity versus no-severe-toxicity groups. We then performed stability selection in combination with least absolute shrinkage and selection operator (LASSO) regression on the set of autoAbs detected at higher levels to identify a parsimonious subset to constitute the final signature (12,13). The prediction score is the linear combination of autoAbs in the signature weighted by their corresponding coefficients in the LASSO regression estimated on the training dataset. These signatures were used to generate prediction scores for patients in the test and independent validation sets.

Statistical analysis

This study conformed to the REMARK guidelines (14). Descriptive group comparisons were performed using two-sample t tests and χ^2 tests for continuous and categorical variables, respectively. The ability of the autoAb signatures to predict recurrence or severe toxicity was first evaluated by determining the AUC of the receiver operating characteristic curve. For each signature, we then identified optimal thresholds from the training sets and used those to classify patients from the test and independent validation sets as having either a high or low autoAb outcome score. In the training dataset, we prioritized cut points that achieved highly accurate predictions of no recurrence (versus recurrence) with a minimum negative predictive value (NPV) of 80% because the decision to not give adjuvant ICB in high-risk resected stage III/IV melanoma requires a high degree of confidence that treatment will not be efficacious. Similarly, we required the cut point for no severe toxicity have a minimum NPV of 80%. Waterfall plots were generated to illustrate the performance of the toxicity classifiers. Two-sample t tests were performed to compare the autoAb toxicity score for patients who did and did not experience severe irAEs. The recurrence-free survival (RFS) of patients with a high versus low autoAb recurrence score was compared by Kaplan–Meier analysis using the log-rank test. The HR with 95% confidence intervals (CI) of patients with a high score relative to those with a low score was estimated by a Cox proportional hazard model. For the patients who received nivolumab, we compared the Cox proportional hazard models with PD-L1 expression (<5% vs. ≥5%) as a single predictor, and with PD-L1 expression (<5% vs. ≥5%) and autoAb recurrence score (high vs. low) both included. Multivariable analyses were performed to estimate the adjusted HR (aHR) of the autoAb score (high vs. low) adjusting the following covariates: age, gender, BRAF mutation status, disease stage, Eastern Cooperative Oncology Group (ECOG) performance status, baseline lactate dehydrogenase (LDH; U/L), and PD-L1

expression (<5% vs. ≥5%). The clinical utility of the autoAb classifier was also assessed by comparing the AUC to that of a model that contained the abovementioned covariates aside from autoAb recurrence score, by Delong test. All statistical analyses were conducted using R version 4.1.0.

To test the potential functional relevance of the identified signatures, we used Metascape to perform broad-based functional analyses of the differentially expressed autoAbs (DEAs) detected at higher levels (https://metascape.org).

Data availability

Data are available from the authors upon request but may require data transfer agreements. No personalized health information will be shared.

Results

Patient characteristics

Table 1 summarizes the baseline clinicodemographic characteristics for all patients included in this study. The clinical and demographic characteristics at baseline were well balanced between training, testing, and independent validation sets for each

treatment regimen (Supplementary Tables S1–S3) and were consistent with the CheckMate 238 and CheckMate 915 total patient populations (1, 10). Supplementary Figures S1 and S2 summarize the recurrence and severe toxicity outcomes for the patients from CheckMate 238 and 915, respectively. Commonly experienced irAEs are summarized in Supplementary Table S4. Notably, most of the severe irAEs were of gastrointestinal, hepatic, or dermatologic etiology; other more serious toxicities were rare in this dataset (i.e., cardiac $n = 3$; diabetes $n = 7$).

Baseline serum autoAb signatures predict recurrence and severe toxicity following adjuvant nivolumab

The nivolumab recurrence signature performed with AUC 0.84 (95% CI, 0.71–0.97) on the test set and AUC 0.82 (95% CI, 0.75–0.88) on the independent validation set. Patients with a high autoAb recurrence score had significantly worse RFS than those with a low score. In the test set, the median RFS was 25.2 months (95% CI, 6.7 months–not reached) for patients with a high recurrence score versus not reached for patients with a low score (RFS monitoring cutoff 44 months), for a HR of 3.71 (95% CI, 1.18–11.67; $P = 0.025$; **Fig. 1A**). In the independent validation set, the median RFS was 16.7 months (95% CI, 12.1–27.1 months) for

Table 1. Baseline clinical and demographic characteristics of patients from each of the four clinical trial cohorts.

	Nivolumab (CheckMate 238)	Nivolumab (CheckMate 915)	Ipilimumab	Ipilimumab plus nivolumab
N	157	190	408	195
Age (years)				
Mean (SD)	55.11 (14.00)	56.39 (12.92)	53.85 (13.35)	56.39 (12.92)
Sex				
Female	68 (43.3%)	80 (42.1%)	169 (41.4%)	89 (45.6%)
Male	89 (56.7%)	110 (57.9%)	239 (58.6%)	106 (54.4%)
Race				
White	153 (97.5%)	189 (99.5%)	391 (95.8%)	191 (97.9%)
Asian	4 (2.5%)	1 (0.5%)	16 (3.9%)	0 (0.0%)
Disease stage[a]				
IIIB	64 (40.8%)	58 (30.5%)	137 (33.6%)	50 (20.6%)
IIIC	60 (38.2%)	97 (51.1%)	190 (46.6%)	111 (56.9%)
IIID	0 (0.0%)	8 (4.2%)	0 (0.0%)	8 (4.1%)
IV	31 (19.7%)	27 (14.2%)	81 (19.9%)	26 (13.3%)
BRAF mutation status				
Wild-type	73 (46.5%)	93 (48.9%)	190 (46.6%)	85 (43.6%)
Mutant	66 (42.0%)	53 (27.9%)	174 (42.6%)	66 (33.8%)
Not reported	18 (11.5%)	44 (23.2%)	44 (10.8%)	44 (22.6%)
Baseline ECOG				
0	140 (89.2%)	183 (96.3%)	363 (89.0%)	181 (92.8%)
1	17 (10.8%)	7 (3.7%)	45 (11.0%)	14 (7.2%)
Baseline LDH (U/L)				
Mean (SD)	225.38 (97.52)	208.69 (64.23)	215.19 (88.36)	223.15 (81.85)
PD-L1 status				
<5%	94 (59.9%)	147 (77.4%)	253 (62.0%)	142 (72.8%)
≥5%	56 (35.7%)	43 (22.6%)	143 (35.0%)	53 (27.2%)
Not reported	7 (4.5%)	0 (0.0%)	12 (2.9%)	0 (0.0%)
Disease recurrence				
No	99 (63.1%)	129 (67.9%)	190 (46.6%)	121 (62.1%)
Yes	58 (36.9%)	61 (32.1%)	218 (53.4%)	74 (37.9%)
Severe toxicity				
No	93 (59.2%)	148 (77.9%)	202 (49.5%)	131 (67.2%)
Yes	64 (40.8%)	42 (22.1%)	206 (50.5%)	64 (32.8%)

[a]Patients in CheckMate 238 were staged using the AJCC Staging Manual 7th Edition. Patients in CheckMate 915 were staged using the AJCC Staging Manual 8th Edition.
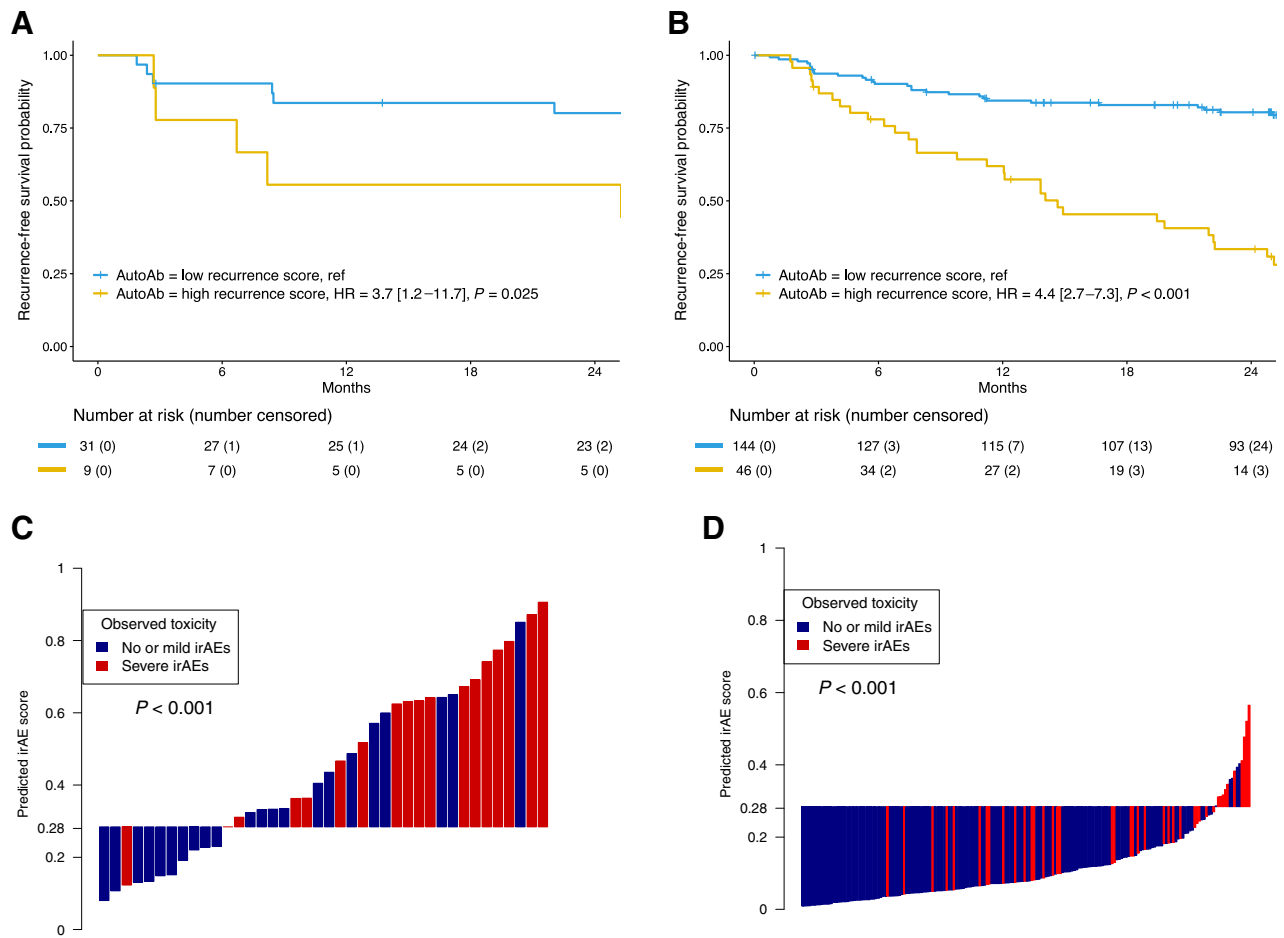
Figure 1.
RFS and severe toxicity by autoAb signature predictions for patients treated with nivolumab. Kaplan–Meier estimates of RFS among patients with high versus low autoAb recurrence scores in the (**A**) test set from CheckMate 238 and (**B**) independent validation set from CheckMate 915. Waterfall plots illustrate the relationship between the predicted and actual development of severe toxicity in patients from the (**C**) test set and (**D**) independent validation set.

patients with a high recurrence score versus not reached for patients with a low score (RFS monitoring cutoff 36 months), for a HR of 4.42 (95% CI, 2.67–7.34; $P < 0.001$; **Fig. 1B**).

The nivolumab severe toxicity signature predicted severe toxicity with AUC 0.78 (95% CI, 0.63–0.93) on the test set and AUC 0.75 (95% CI, 0.67–0.83) on the independent validation set. In both the test and independent validation sets, the autoAb severe toxicity scores were significantly higher in patients who experienced severe irAEs ($P < 0.001$ for each; **Fig. 1C** and **D**). In multivariable analyses, autoAb severe toxicity score remained a significant predictor of severe toxicity (aHR, 13.53; 95% CI, 2.59–86.65; $P = 0.003$; Supplementary Table S5). The autoAb signatures are listed in Supplemental File 1.

The nivolumab recurrence and severe toxicity signatures could be used together to stratify patients into four possible outcomes combinations for the test (**Fig. 2A**) and independent validation sets (**Fig. 2B**). In the independent validation set, 132 patients were predicted on the basis of the composite signature to have no recurrence and no severe toxicity; of these, 103 (78%) did not recur, and 23 (17%) developed severe toxicity. Forty-two patients were predicted to suffer recurrence but no severe toxicity; 28 (67%) recurred and 8 (19%) had severe irAEs. Twelve patients were predicted to have no recurrence but severe toxicity; of these, only 1 (8%) recurred and 7 (58%) had severe toxicity.

Only 4 patients were predicted to have recurrence and severe toxicity; 4 (100%) had severe toxicity and 3 (75%) recurred.

Nivolumab autoAb recurrence signatures outperform clinical variables, including PD-L1 status

In the nivolumab test set, there was no significant difference in RFS between patients with PD-L1 <5% versus ≥5% (HR, 2.75; 95% CI, 0.59–12.79; $P = 0.197$; **Fig. 3A**). In the independent validation set, PD-L1 ≥ 5% was associated with significantly better RFS (HR, 2.71; 95% CI, 1.40–5.23; $P = 0.003$; **Fig. 3B**). When analyzed jointly, high autoAb recurrence score (aHR, 3.60; 95% CI, 1.98–6.55; $P < 0.001$) and PD-L1 <5% (aHR, 2.10; 95% CI, 1.02–4.30; $P = 0.043$) were significant predictors of worse RFS (Supplementary Table S6). In the test and independent validation sets, patients with PD-L1 <5% and a high recurrence score had the worst RFS (test: HR, 8.43; 95% CI, 0.93–76.05; $P = 0.058$; independent validation: HR, 10.05; 95% CI, 4.08–24.73; $P < 0.001$; **Fig. 3C** and **D**). Their RFS was worse than patients with an unfavorable prediction from any single predictor alone, which underscored the added utility of the autoAb recurrence signature prediction. The AUC of the autoAb recurrence signature was better than that of a clinical model that included PD-L1 status (AUC 0.839 vs. 0.663;
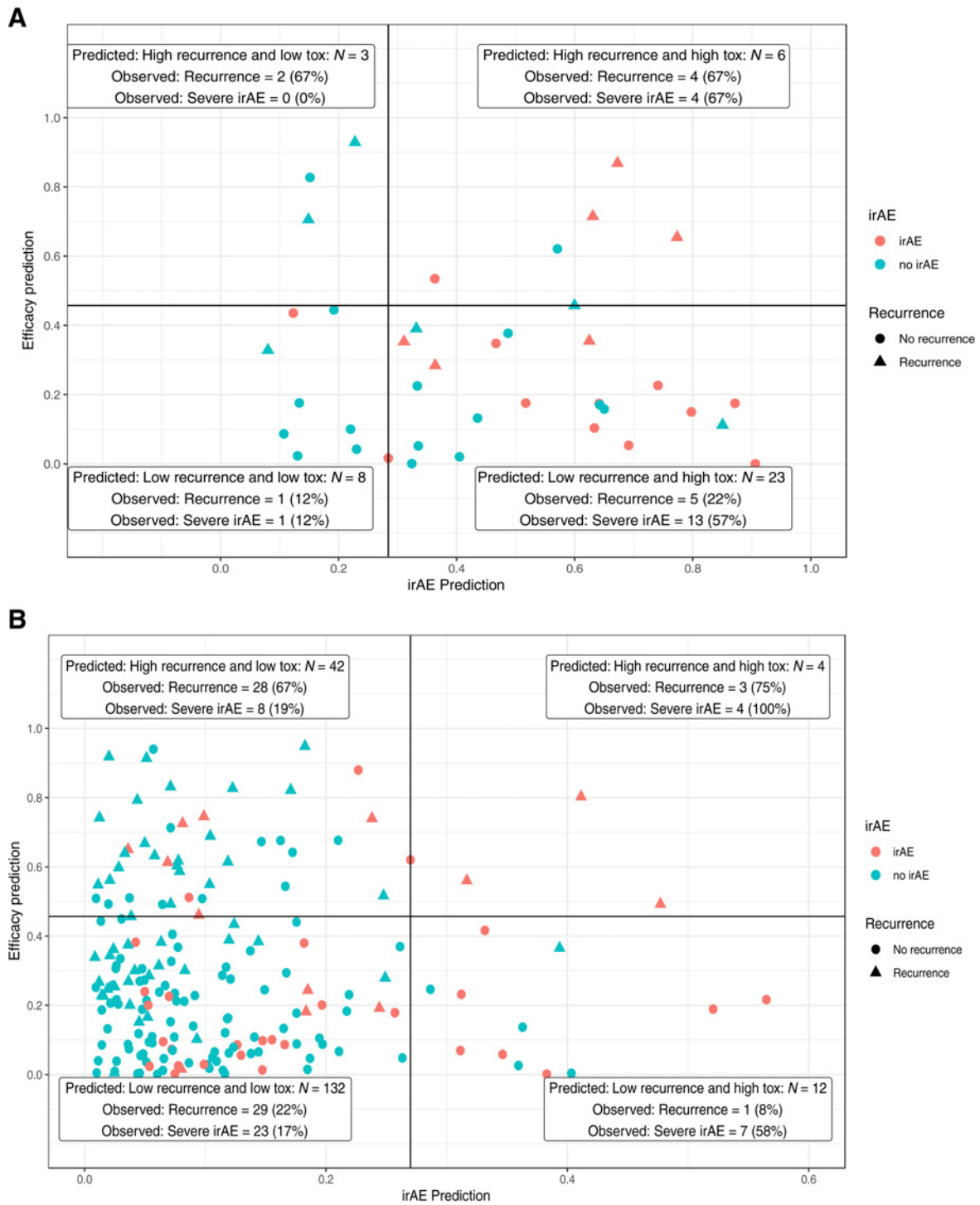
Figure 2.
Stratification of patients treated with nivolumab based on their predicted risk of disease recurrence and severe toxicity. Patients in the nivolumab (**A**) test and (**B**) independent validation sets were stratified into four quadrants based on their projected risk of recurrence and severe toxicity. The quadrants were divided at the cutoffs for the irAE prediction score (*x*-axis) and the recurrence prediction score (*y*-axis) as determined in the training set. Each point represents a distinct patient. The colors indicate the observed severe toxicity outcomes (red, patients who experienced a severe irAE; blue, patients who did not experience a severe irAE) and shapes show the observed recurrence outcomes (circles, patients who did not have recurrence; triangles, patients who did have recurrence).
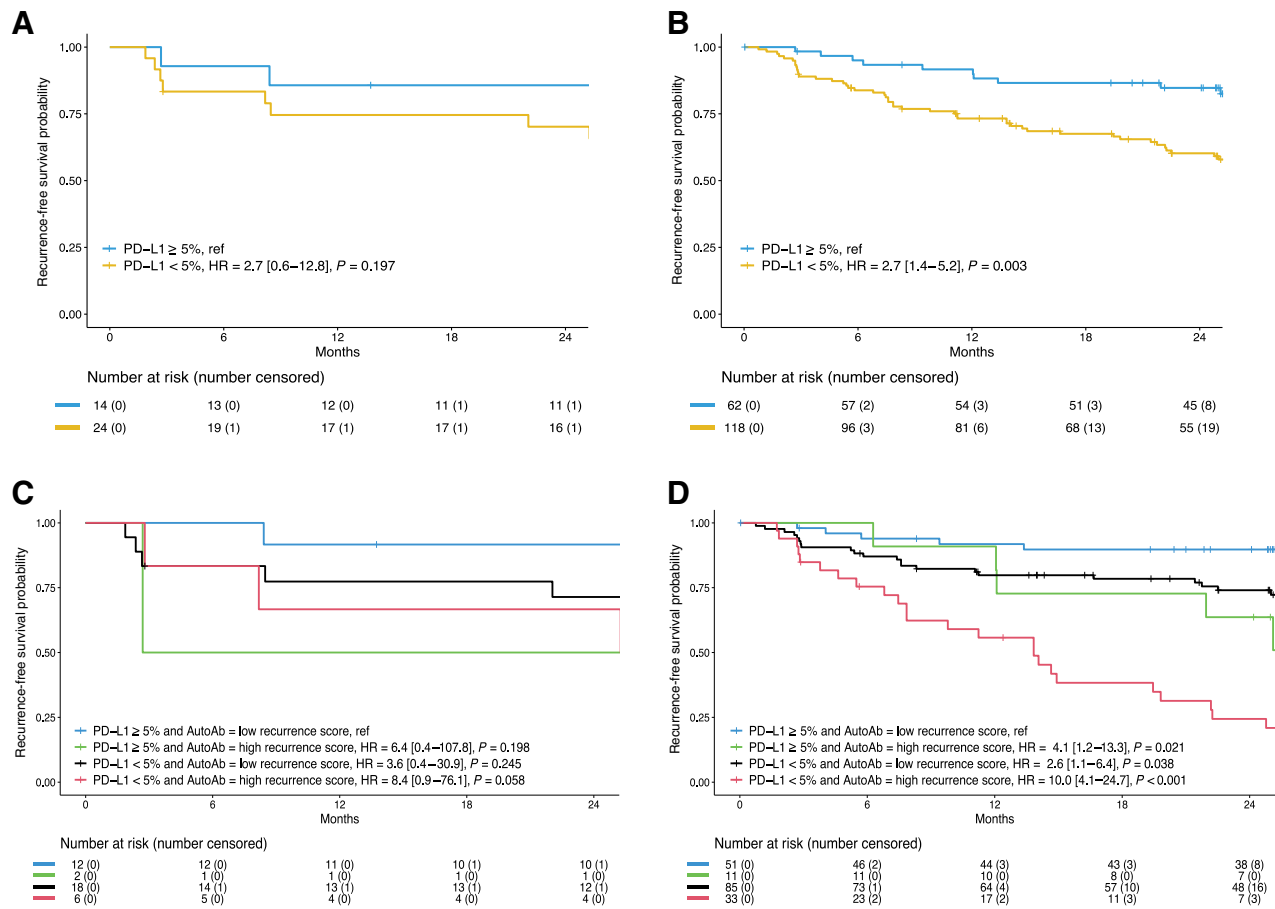
Figure 3.
RFS among patients treated with nivolumab and according to PD-L1 status and PD-L1 status plus autoAb signature prediction. Kaplan–Meier estimates of RFS among patients with PD-L1 expression <5% versus ≥5% in the (**A**) test set from CheckMate 238 and (**B**) independent validation set from CheckMate 915. Kaplan–Meier estimates of RFS were plotted for patients from the (**C**) test and (**D**) independent validation sets based on their PD-L1 expression (<5% vs. ≥5%) and autoAb recurrence score (high vs. low).

$P = 0.12$ in the test set; AUC 0.816 vs. 0.556; $P < 0.001$ in the independent validation set).

Serum autoAbs predict recurrence and severe toxicity in patients treated with ipilimumab or ipilimumab plus nivolumab

The ipilimumab recurrence signature performed with AUC 0.76 (95% CI, 0.66–0.85) on the test set. Patients with a high autoAb recurrence score had significantly worse RFS than those with a low score. In the test set, the median RFS was 15.1 months (95% CI, 7.2 months–not reached) for patients with a high score versus not reached for patients with a low score (RFS monitoring cutoff 44 months), for a HR of 3.19 (95% CI, 1.42–7.15; $P = 0.005$; **Fig. 4A**). In multivariable analyses, high autoAb recurrence score was a significant predictor of RFS (aHR, 3.21; 95% CI, 1.38–7.45; $P = 0.007$; Supplementary Table S6).

The ipilimumab severe toxicity signature achieved AUC 0.79 (95% CI, 0.70–0.88) on the test set. Patients who developed severe toxicity had significantly higher autoAb toxicity scores compared with those who did not ($P < 0.001$; **Fig. 4C**). In multivariable analyses, autoAb toxicity score remained a significant predictor of severe toxicity (aHR, 11.04; 95% CI, 3.84–37.25; $P < 0.001$; Supplementary Table S5). The ipilimumab recurrence and toxicity signatures

could be used together to accurately stratify patients into one of the four different possible outcomes combinations (Supplementary Fig. S3A). The autoAb signatures are listed in Supplemental File 1.

The recurrence signature for ipilimumab plus nivolumab performed with AUC 0.92 (95% CI, 0.85–0.99) on the test set. Patients with a high autoAb recurrence score had significantly worse RFS than those with a low score. In the test set, the median RFS was 17.1 months (95% CI, 13.9 months–not reached) for patients with a high score versus not reached for patients with a low score (RFS monitoring cutoff 36 months), for a HR of 6.37 (95% CI, 2.37–17.10; $P < 0.001$; **Fig. 4B**). AutoAb score remained a significant predictor of RFS in multivariable analyses (aHR, 6.45; 95% CI, 1.48–28.02; $P = 0.013$; Supplementary Table S6).

The ipilimumab plus nivolumab toxicity signature achieved AUC 0.87 (95% CI, 0.75–0.99) on the test set. The patients who developed severe toxicity had significantly higher autoAb toxicity scores compared with patients who did not experience severe toxicity ($P < 0.001$; **Fig. 4D**). AutoAb score remained a significant predictor of severe toxicity in multivariable analyses (aHR, 23.44; 95% CI, 4.10–212.50; $P = 0.001$; Supplementary Table S5). The ipilimumab plus nivolumab recurrence and severe toxicity signatures could be used together to accurately stratify patients
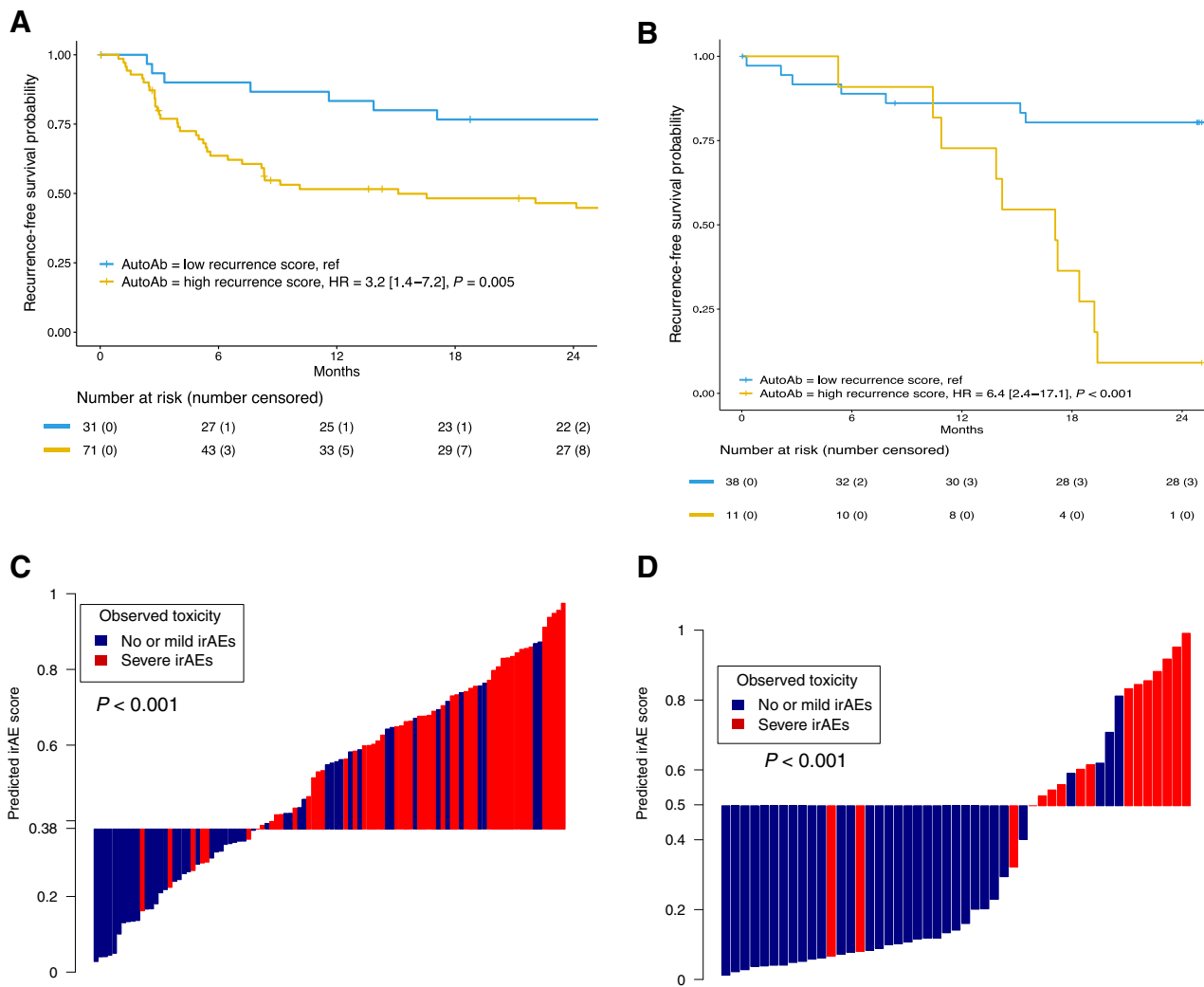
Figure 4.
RFS and severe toxicity by autoAb signature predictions, evaluated in the ipilimumab and the ipilimumab plus nivolumab test sets. Kaplan–Meier estimates of RFS in (**A**) ipilimumab patients with high versus low autoAb recurrence scores, and (**B**) ipilimumab plus nivolumab patients with high versus low autoAb recurrence scores. Waterfall plots show the relationship between the predicted and actual development of severe toxicity for (**C**) ipilimumab and (**D**) ipilimumab plus nivolumab patients.

into four different possible outcomes combinations (Supplementary Fig. S3B). The autoAb signatures are listed in Supplemental File 1.

Functional and enrichment analyses

Patients' median overall autoAb signal intensity was not associated with disease recurrence or severe toxicity for any of the three treatment regimens ($P > 0.20$ for all). There was no significant association between disease recurrence and the development of severe toxicity for any treatment regimen ($P > 0.50$ for all comparisons). There was a statistically significant overlap between the DEAs associated with recurrence and severe toxicity for ipilimumab monotherapy as well as for ipilimumab plus nivolumab, but not for the DEAs associated with nivolumab recurrence and severe toxicity (Supplementary Fig. S4). There was no significant overlap of autoAb between the final signatures (Supplementary Fig. S5). The autoAbs detected at higher levels associated with disease recurrence for all

three regimens were enriched for antigens related to the "negative regulation of immune system process," and both the nivolumab and ipilimumab profiles were enriched for "inflammatory response" (Supplementary Fig. S6B). The analysis of DEA associated with severe toxicity showed overlapping enrichment for antigens involved with the immune-related pathways "inflammatory response" and "chemotaxis" (Supplementary Fig. S6C).

Discussion

We identified and validated a composite panel of pretreatment serum autoAbs that predicted recurrence and severe toxicity in melanoma patients who received adjuvant nivolumab. These autoAb signatures could be used together to stratify patients on the basis of their projected likelihood of suffering relapsed disease and developing severe irAEs. This is in contrast with most biomarkers under active investigation, which aim to predict either treatment efficacy or toxicity.

The ability to simultaneously forecast both outcomes would enable providers and patients to assess the possibility of clinical benefit in the context of potential toxicity and ultimately help optimize treatment regimens while minimizing exposure to severe irAEs.

We also identified serum autoAb signatures that predicted recurrence and severe toxicity for patients who receive adjuvant ipilimumab or ipilimumab plus nivolumab. These analyses were limited by the lack of independent validation cohorts because each regimen was included as an arm in one of the phase III trials but not the other, and because they are not commonly used as standard of care in the adjuvant setting. In contrast, there were randomized nivolumab arms in both CheckMate 238 and CheckMate 915, which allowed for independent validation of the identified autoAb signatures through prospective-specimen-collection and retrospective-blinded-evaluation, which is a recommended approach for pivotal evaluation of the accuracy of a biomarker used for prediction (15). Nevertheless, the promising results obtained in the test cohorts for ipilimumab and ipilimumab plus nivolumab suggest the potential for identifying the subset of patients most likely to benefit from either of those therapies in the adjuvant setting.

Pretreatment tumor cell expression of PD-L1 was one of the earliest and most extensively studied candidate biomarkers of ICB response. Although some data suggest that higher PD-L1 expression is associated with better treatment outcomes in different cancers, PD-L1 status has limited predictive utility and many patients classified as negative still respond to anti–PD-1 therapy (16–18). We found that the parsimonious autoAb recurrence signature for nivolumab predicted treatment outcomes more accurately than a multivariable model composed of clinical covariates including PD-L1 status. The autoAb signatures identified in this study offer other advantages. From a technical standpoint, the autoAb readouts and the signature algorithms can be applied to each patient using their own internal control, which is the median overall signal intensity on that individual's panel. This obviates the need to use external controls or normalizing processes when testing serum from new patients.

We focused our investigation on the predictive potential of antibodies detected at higher levels in patients without recurrence or with severe toxicity. Our decision to do so was predicated on our *a priori* hypothesis that the presence of autoAbs indicates subclinical immune system activity, which itself foreshadows the immune system activation seen in patients who respond to treatment or develop severe immunotoxicity. We did not limit the signature building process to autoAbs whose level exceeded a certain number of SDs above that of the healthy controls, as is often done in autoimmune disease antibody research (19, 20). While our findings illustrate the predictive value of focusing on subclinical but higher-level antibodies, we cannot exclude the possibility that antibodies detected at lower levels might also contribute biologically to the likelihood of developing severe toxicity or not suffering disease recurrence.

We noted minimal overlap between the autoAbs detected at higher levels associated with severe toxicity and disease recurrence, which suggests that different underlying mechanisms might drive ICB toxicity and therapeutic effect. While there is a growing body of evidence that the occurrence of irAEs portends better immunotherapy response in patients with nonmelanoma solid tumors, the correlation between efficacy and toxicity does not appear to be as evident for individuals with melanoma (21, 22). Moreover, retrospective analyses of both melanoma and nonmelanoma patients showed that the survival benefits are limited to individuals who experience low- but not high-grade toxicity (23–25). This may be due to a constellation of factors including protracted treatment interruptions, the use of high-dose steroids or other potent immunosuppressive agents, and significant end-organ damage for patients with severe irAEs.

There is increasing recognition that the B-cell compartment impacts ICB response, but the role of antibody-mediated immunity has not been fully elucidated (26, 27). In one recent study, higher levels of total IgG at baseline correlated with better progression-free survival in patients with metastatic melanoma treated with ICB (28). In contrast, we observed that the median overall antibody signal intensity was not associated with disease recurrence. Our findings indirectly raise the question whether the immunogenicity of specific antigens might predict ICB outcomes more effectively than the level of total antibody synthesis. Das and colleagues (2018) reported on the impact of the B-cell compartment on toxicity outcomes by showing that early changes in B cells correlated with higher rates of severe toxicity (29). Several other studies demonstrated that there is overlap between antibodies associated with organ-specific autoimmune diseases and irAEs specific to those same organs. For instance, baseline levels of antithyroid antibodies correlate with the future development of immunotherapy-related thyroid toxicity (6). The primary goal of checkpoint blockade is to activate cytotoxic T cells and there are extensive data to implicate T-cell reactivity in the development of irAEs (30–32). Robert and colleagues (2014) showed that CTLA-4 inhibition is associated with peripheral T-cell receptor expansion and posited that this could indicate the mobilization of autoreactive T cells (33). In a similar vein, Oh and colleagues (2017) reported that CTLA-4 blockade led to greater diversification of the T-cell repertoire in patients who developed irAEs compared with those who did not (34). When taken together, these data suggest that subclinical autoimmunity might play a role in the risk of developing irAEs.

Strengths and limitations

Our study has multiple strengths. First, the datasets come from two phase III studies so the serum collection, treatment schedules, and outcomes assessments were all standardized and well annotated. Second, the autoAb signatures outperformed a multivariable clinical model that includes percent PD-L1 expression, which is the current clinical benchmark for identifying patients likely to respond to immunotherapy. Third, we were able to independently validate the nivolumab recurrence and severe toxicity signatures and show that both perform consistently across multiple doses and treatment schedules from 2 different clinical trials.

There are also some limitations to this study. The patients enrolled in CheckMate 238 and CheckMate 915 were treatment naïve. Growing evidence suggests that the humoral immune system evolves after ICB, which might affect outcomes in subsequent lines of immunotherapy (35). Although the data were prospectively acquired in the phase III clinical trial setting, our study was nevertheless retrospective so the identified autoAb signatures require further evaluation in the prospective setting. Future testing on additional cohorts will be needed to verify the generalizability of the models. In addition, the predictive utility of the signatures will need to be validated across multiple different treatment doses and schedules. The signatures will also need to be tested on other autoAb microarrays in addition to the platform provided by CDI Laboratories. Finally, the most common severe adverse toxicities in our cohorts included gastrointestinal, hepatic, and dermatologic events. Going forward, it will be important to test the applicability of the severe toxicity signatures on less common organ-specific toxicities such as encephalitis, meningitis, myocarditis, pericarditis, and type 1 diabetes mellitus. However, we could not perform this analysis because the incidence of these events was low in the available cohorts. Nevertheless, we believe that the ability to predict

any severe toxicity has the potential to impact management in several ways. For instance, if a patient is predicted to be at high risk of recurrence on nivolumab, but not ipilimumab, and they are predicted to experience severe toxicity on both, the decision could be made in favor of ipilimumab considering the durable clinical benefit that some patients experience with CTLA-4 blockade. Conversely, the autoAb signatures could help minimize exposure to immune-related toxicity if two regimens are forecasted to be efficacious but only one is expected to lead to severe toxicity. This could in turn help augment survival outcomes because the management of severe irAEs often includes systemic immunosuppression, which can affect the adaptive antitumor immune response and in turn negatively impact immunotherapy efficacy (25). We acknowledge, however, that this would need to be tested in biomarker-driven clinical trial.

Conclusions

The panel of baseline autoAbs identified in this study predicted whether resected melanoma patients were at risk for disease recurrence or severe toxicity after treatment with adjuvant ICB. It will be imperative to explore whether on-treatment autoAb profiles predict immunotherapy treatment outcomes and to what degree they correspond with pretreatment signatures. This could offer a means for serial monitoring of treatment progress. Finally, it will be interesting to test the predictive utility of these autoAb signatures in patients with the multiple other cancers for which immune checkpoint inhibitors received FDA approval.

Authors' Disclosures

D. Fenyo reports personal fees from Spectragen Informatics and Proteome Software and other support from Preverna, Protein Metrics, and The Informatics Factory outside the submitted work. M. Wind-Rotolo reports other support from Bristol-Myers Squibb during the conduct of the study as well as other support from Agios Pharmaceuticals outside the submitted work. M. Krogsgaard reports personal fees from Neximmune, Guidepoint, and Repertoire; grants and personal fees from Genentech/Roche; and grants from Novartis, Merck, and Mark Foundation for Cancer Research outside the submitted work. J.M. Mehnert reports grants and personal fees from Bristol-Myers Squibb, Novartis, and Regeneron and personal fees from Eisai, Merck, and Seagen outside the submitted work; in addition, J.M.

Mehnert has served on advisory boards for Bristol-Myers Squibb, Eisai, Regeneron, Seagen, and Novartis and has served as a consultant for Merck. J.S. Weber reports personal fees from Bristol-Myers Squibb during the conduct of the study; in addition, J.S. Weber has a patent for a PD-1 biomarker from Biodesix issued. I. Osman reports grants from NCI during the conduct of the study; in addition, I. Osman has a patent pending to AutoAB signatures to predict immunotherapy response and toxicity, which has been submitted and reported through iEdison, Invention Report Number 5998301-21-0004. No disclosures were reported by the other authors.

Authors' Contributions

P. Johannet: Conceptualization, data curation, formal analysis, validation, investigation, methodology, writing–original draft, writing–review and editing. **W. Liu:** Formal analysis, investigation, writing–review and editing. **D. Fenyo:** Investigation, writing–review and editing. **M. Wind-Rotolo:** Resources. **M. Krogsgaard:** Investigation, writing–review and editing. **J.M. Mehnert:** Investigation, writing–review and editing. **J.S. Weber:** Conceptualization, resources, funding acquisition, investigation, writing–review and editing. **J. Zhong:** Conceptualization, data curation, software, formal analysis, supervision, funding acquisition, validation, investigation, visualization, methodology, writing–original draft, project administration, writing–review and editing. **I. Osman:** Conceptualization, resources, data curation, formal analysis, supervision, funding acquisition, validation, investigation, visualization, methodology, writing–original draft, project administration, writing–review and editing.

Acknowledgments

Note

Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

References

1. Weber J, Mandala M, Del Vecchio M, Gogas HJ, Arance AM, Cowey CL., et al. Adjuvant nivolumab versus ipilimumab in resected stage III or IV MELANO-MA. N Engl J Med 2017;377:1824–35.
2. Eggermont AMM, Chiarion-Sileni V, Grob J-J, Dummer R, Wolchok JD, Schmidt H, et al. Prolonged survival in stage III melanoma with ipilimumab adjuvant therapy. N Engl J Med 2016;375:1845–55.
3. Postow MA, Sidlow R, Hellmann MD. Immune-related adverse events associated with immune checkpoint blockade. N Engl J Med 2018;378:158–68.
4. Michot JM, Bigenwald C, Champiat S, Collins M, Carbonnel F, Postel-Vinay S, et al. Immune-related adverse events with immune checkpoint blockade: a comprehensive review. Eur J Cancer 2016;54:139–48.
5. Tahir SA, Gao J, Miura Y, Blando J, Tidwell RSS, Zhao H, et al. Autoimmune antibodies correlate with immune checkpoint therapy-induced toxicities. Proc Natl Acad Sci USA 2019;116:22246–51.
6. Kimbara S, Fujiwara Y, Iwama S, Ohashi K, Kuchiba A, Arima H, et al. Association of antithyroglobulin antibodies with the development of thyroid dysfunction induced by nivolumab. Cancer Sci 2018;109:3583–90.
7. Gowen MF, Giles KM, Simpson D, Tchack J, Zhou H, Moran U, et al. Baseline antibody profiles predict toxicity in melanoma patients treated with immune checkpoint inhibitors. J Transl Med 2018;16:82.
8. Fässler M, Diem S, Mangana J, Hasan Ali O, Berner F, Bomze D, et al. Antibodies as biomarker candidates for response and survival to checkpoint inhibitors in melanoma patients. J Immunother Cancer 2019;7:50.

9. Toi Y, Sugawara S, Sugisaka J, Ono H, Kawashima Y, Aiba T, et al. Profiling preexisting antibodies in patients treated with anti–PD-1 therapy for advanced non–small cell lung cancer. JAMA Oncol 2019;5:376–83.
10. Long GV, Schadendorf D, Del Vecchio M, Larkin J, Atkinson V, Schenker M, et al. Adjuvant therapy with nivolumab (NIVO) combined with ipilimumab (IPI) vs NIVO alone in patients (pts) with resected stage IIIB-D/IV melanoma (CheckMate 915) [abstract]. In: Proceedings of the American Association of Cancer Research Annual Meeting 2021; 2021 Apr 10–15 and May 17–21. Philadelphia (PA): AACR, Cancer Res 2021; 81 (13_Suppl);Abstract nr CT004.
11. Jeong JS, Jiang L, Albino E, Marrero J, Rho HS, Hu J, et al. Rapid identification of monospecific monoclonal antibodies using a human proteome microarray. Mol Cell Proteomics 2012;11:O111.016253.
12. Tibshirani R. Regression shrinkage and selection via the lasso. J Royal Stat Soc: Series B (Statistical Methodology), 1996;58:267–88.
13. Meinshausen N, Buhlmann P. Stability selection. J Royal Stat Soc: Series B (Statistical Methodology) 2010;72:417–73.
14. Altman DG, Mcshane LM, Sauerbrei W, Taube SE, Reporting recommendations for tumor marker prognostic studies (REMARK): explanation and elaboration. BMC Med 2012;10:51.
15. Pepe MS, Feng Z, Janes H, Bossuyt PM, Potter JD. Pivotal evaluation of the accuracy of a biomarker used for classification or prediction: standards for study design. J Natl Cancer Inst 2008;100:1432–8.

16. Daud AI, Wolchok JD, Robert C, Hwu W-J, Weber JS, Ribas A, et al. Programmed death-ligand 1 expression and response to the anti-programmed death 1 antibody pembrolizumab in melanoma. J Clin Oncol 2016;34:4102–9.

17. Gibney GT, Weiner LM, Atkins MB. Predictive biomarkers for checkpoint inhibitor-based immunotherapy. Lancet Oncol 2016;17:e542–51.

18. Topalian SL, Hodi FS, Brahmer JR, Gettinger SN, Smith DC, Mcdermott DF, et al. Safety, activity, and immune correlates of anti–PD-1 antibody in cancer. N Engl J Med 2012;366:2443–54.

19. De Moel EC, Derksen VFAM, Trouw LA, Bang H, Collée G, Lard LR, et al. In rheumatoid arthritis, changes in autoantibody levels reflect intensity of immunosuppression, not subsequent treatment response. Arthritis Res Ther 2019;21:28.

20. Pardos-Gea J, Cortés-Hernández J, Castro-Marrero J, Balada E, Ordi-Ros J. Autoantibodies to types I and IV collagen and heart valve disease in systemic lupus erythematosus/antiphospholipid syndrome. Clin Rheumatol 2017;36: 1401–6.

21. Das S, Johnson DB. Immune-related adverse events and antitumor efficacy of immune checkpoint inhibitors. J Immunother Cancer 2019;7:306.

22. Weber JS, Hodi FS, Wolchok JD, Topalian SL, Schadendorf D, Larkin J, et al. Safety profile of nivolumab monotherapy: a pooled analysis of patients with advanced melanoma. J Clin Oncol 2017;35:785–92.

23. Matsuoka H, Hayashi T, Takigami K, Imaizumi K, Shiroki R, Ohmiya N, et al. Correlation between immune-related adverse events and prognosis in patients with various cancers treated with anti–PD-1 antibody. BMC Cancer 2020;20:656.

24. Zhou X, Yao Z, Yang H, Liang N, Zhang X, Zhang F. Are immune-related adverse events associated with the efficacy of immune checkpoint inhibitors in patients with cancer? A systematic review and meta-analysis. BMC Med 2020;18:87.

25. Eggermont AMM, Kicinski M, Blank CU, Mandala M, Long GV, Atkinson V, et al. Association between immune-related adverse events and recurrence-free survival among patients with stage III melanoma randomized to receive pembrolizumab or placebo: a secondary analysis of a randomized clinical trial. JAMA Oncol 2020;6:519–27.

26. Helmink BA, Reddy SM, Gao J, Zhang S, Basar R, Thakur R, et al. B cells and tertiary lymphoid structures promote immunotherapy response. Nature 2020; 577:549–55.

27. Griss J, Bauer W, Wagner C, Simon M, Chen M, Grabmeier-Pfistershammer K, et al. B cells sustain inflammation and predict response to immune checkpoint blockade in human melanoma. Nat Commun 2019;10:4186.

28. Diem S, Fässler M, Bomze D, Ali OH, Berner F, Niederer R, et al. Immunoglobulin G and subclasses as potential biomarkers in metastatic melanoma patients starting checkpoint inhibitor treatment. J Immunother 2019;42:89–93.

29. Das R, Bar N, Ferreira M, Newman AM, Zhang L, Bailur JK, et al. Early B cell changes predict autoimmunity following combination immune checkpoint blockade. J Clin Invest 2018;128:715–20.

30. Lozano AX, Chaudhuri AA, Nene A, Bacchiocchi A, Earland N, Vesely MD, et al. T cell characteristics associated with toxicity to immune checkpoint blockade in patients with melanoma. Nat Med 2022;28:353–62.

31. Luoma AM, Suo S, Williams HL, Sharova T, Sullivan K, Manos M, et al. Molecular pathways of colon inflammation induced by cancer immunotherapy. Cell 2020;182:655–71.

32. Sasson SC, Slevin SM, Cheung VTF, Nassiri I, Olsson-Brown A, Fryer E, et al. Interferon-gamma-producing CD8$^+$ tissue resident memory t cells are a targetable hallmark of immune checkpoint inhibitor-colitis. Gastroenterology 2021;161:1229–44.

33. Robert L, Tsoi J, Wang X, Emerson R, Homet B, Chodon T, et al. CTLA4 blockade broadens the peripheral T-cell receptor repertoire. Clin Cancer Res 2014;20:2424–32.

34. Oh DY, Cham J, Zhang L, Fong G, Kwek SS, Klinger M, et al. Immune toxicities elicted by CTLA-4 blockade in cancer patients are associated with early diversification of the T-cell repertoire. Cancer Res 2017;77:1322–30.

35. De Moel EC, Rozeman EA, Kapiteijn EH, Verdegaal EME, Grummels A, Bakker JA, et al. Autoantibody development under treatment with immune checkpoint inhibitors. Cancer Immunol Res 2019;7:6–11.