Journal for ImmunoTherapy of Cancer

Early dynamics of clinical and laboratory parameters predict primary refractory disease in patients with metastatic urothelial carcinoma receiving atezolizumab

Christopher J Graser ,[1,2,3] Thomas O McDonald ,[1,2,3,4] Paul J Catalano ,[1,2] Guru Sonpavde ,[5,6] Franziska Michor [1,2,3,4,7,8]

To cite: Graser CJ, McDonald TO, Catalano PJ, *et al*. Early dynamics of clinical and laboratory parameters predict primary refractory disease in patients with metastatic urothelial carcinoma receiving atezolizumab. *Journal for ImmunoTherapy of Cancer* 2025;**13**:e011740. doi:10.1136/jitc-2025-011740

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/jitc-2025-011740).

Accepted 31 March 2025

For numbered affiliations see end of article.

Correspondence to
Dr Franziska Michor; michor@jimmy.harvard.edu

Dr Guru Sonpavde; guru.sonpavde.md@adventhealth.com

ABSTRACT

Background In patients with metastatic urothelial carcinoma (mUC) receiving programmed cell death ligand 1 (PD-L1) inhibitors, it is critically important to identify primary refractory patients very early to enable modification of therapy before clinical progression and decline of performance status. We hypothesized that baseline and early-on-treatment (EOT) parameters may help identify patients likely to have primary refractory disease.

Methods We considered baseline and EOT variables measured up to 5 weeks after initiating therapy in the phase 3 clinical trial IMvigor211, which compared atezolizumab versus chemotherapy, in muC patients who had progressed on platinum-based chemotherapy. We used least absolute shrinkage and selection operator-regularized logistic regression models to predict the risk of primary refractory disease employing clinical and laboratory variables.

Results 902 patients were evaluable for analysis. Our baseline model achieves an area under the curve (AUC) of 0.730, 0.717 for the atezolizumab group and 0.696 for the chemotherapy group. The AUC increases to 0.848 overall with EOT parameters, 0.871 for the atezolizumab group and 0.788 for the chemotherapy group. The EOT model suggests that 33.7% of patients receiving atezolizumab may benefit from switching to chemotherapy, reducing their risk of primary refractoriness from 67.1% to 51.5%.

Conclusions Our prediction model employs readily available and routinely measured clinical and laboratory factors, such as urine-specific gravity, presence of liver metastases, and total protein and erythrocyte counts. It robustly identifies patients with early primary refractory disease to atezolizumab before clinical progression and may inform therapeutic decisions. Validation in larger independent cohorts and other treatments is required.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Treatment outcomes with PD1/L1 inhibitors exhibit substantial variability among patients with metastatic urothelial carcinoma, with approximately 20% of patients showing durable responses, while around 35% experience primary refractory disease. Early identification of the primary refractory disease population is important to avoid futile and toxic therapy.

WHAT THIS STUDY ADDS

⇒ By integrating baseline and early-on-treatment information, our approach improves on current prediction models that employ baseline factors. Moreover, we identified several clinical features differentially predictive of treatment outcomes for both atezolizumab and chemotherapy.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Our model is affordable, relying on routine clinical and laboratory variables, and may inform an early switch in therapy. Our findings constitute a proof of principle that early dynamics in such variables hold predictive value for absolute risks of refractory disease, as well as differences in risk among treatments. Our approach can easily be extended to other immunotherapy settings.

INTRODUCTION

Patients with metastatic urothelial carcinoma (mUC) who are treated with PD1/L1 inhibitors exhibit substantial variability in outcomes. Objective response rates for patients receiving PD1/L1 inhibitors for progressive disease (PD) after platinum-based chemotherapy range between 15% and 25%, and a large proportion of patients (~30%–40%) exhibit primary refractory disease with PD as the best response.[1–3] While the toxicities of PD1/L1 inhibitors are generally manageable, 15%–20% of patients experience severe side effects, and a fraction of patients suffers from fatal or permanent life-altering events.[1] Hence, it is essential to pre-emptively identify patients at high risk of refractory disease

likely to have early PD to avoid futile therapy and enable an early switch to other treatment choices before early death, rapid clinical progression, or potential ineligibility for second-line therapy due to frailty or decline in performance status.

Predictive models of response have identified some associations between treatment outcomes and clinical variables such as the percentage of PD-L1-positive tumor and/or immune cells,[4–7] tumor mutation burden,[8] and a composite panel of baseline clinical and laboratory factors.[9 10] Indeed, three PD1/L1 inhibitors (pembrolizumab, nivolumab, and avelumab) have been approved as monotherapy for progressive mUC following platinum-based chemotherapy regardless of the levels of PD-L1 expression. Moreover, previous work has focused mostly on identifying prognostic variables at baseline such as ECOG (Eastern Cooperative Oncology Group) performance status, the presence of liver metastases, platelet count, neutrophil-to-lymphocyte ratio, and lactate dehydrogenase at baseline.[11 12] However, little consideration has been given to the dynamics of variables measured after the start of treatment.

We set out to investigate the ability of routine baseline clinical and laboratory measurements combined with early-on-treatment (EOT) variables to improve prediction models of absolutely refractory disease, that is, PD as the best outcome. We estimated a logistic regression model using a large prospective trial dataset to predict treatment outcomes for patients receiving the PD-L1 inhibitor atezolizumab or chemotherapy. This model relies on a large number of covariates, including most prominently urine-specific gravity, presence of liver metastases, T-stage of the primary tumor, total protein, platelet and erythrocyte counts, as well as changes in erythrocyte counts during the first cycle of treatment.

Our model significantly outperforms approaches based on known predictors of treatment success for PD-L1 inhibitors.[11 12] Additionally, our model simultaneously predicts outcomes for patients receiving chemotherapy, identifying candidates for whom this alternative therapy promises better treatment outcomes.

METHODS

Data

We obtained data from the phase 3 clinical trial IMvigor211, which compared the PD-L1 inhibitor atezolizumab vs chemotherapy (taxane or vinflunine) in patients with mUC who had previously progressed on platinum-based combination chemotherapy (cisplatin and/or carboplatin) (figure 1).[13 14] The trial included 931 patients who were randomized to the atezolizumab (n=467) or chemotherapy (n=464) arms. Patients were treated in successive 3-week cycles with treatment administered intravenously on the first day of each cycle. In both groups, treatment was discontinued if a patient showed unacceptable toxicity or the patient withdrew consent. Patients in the chemotherapy group were treated until disease progression, while patients in the atezolizumab arm were treated until "symptomatic deterioration attributed to disease progression".[13]

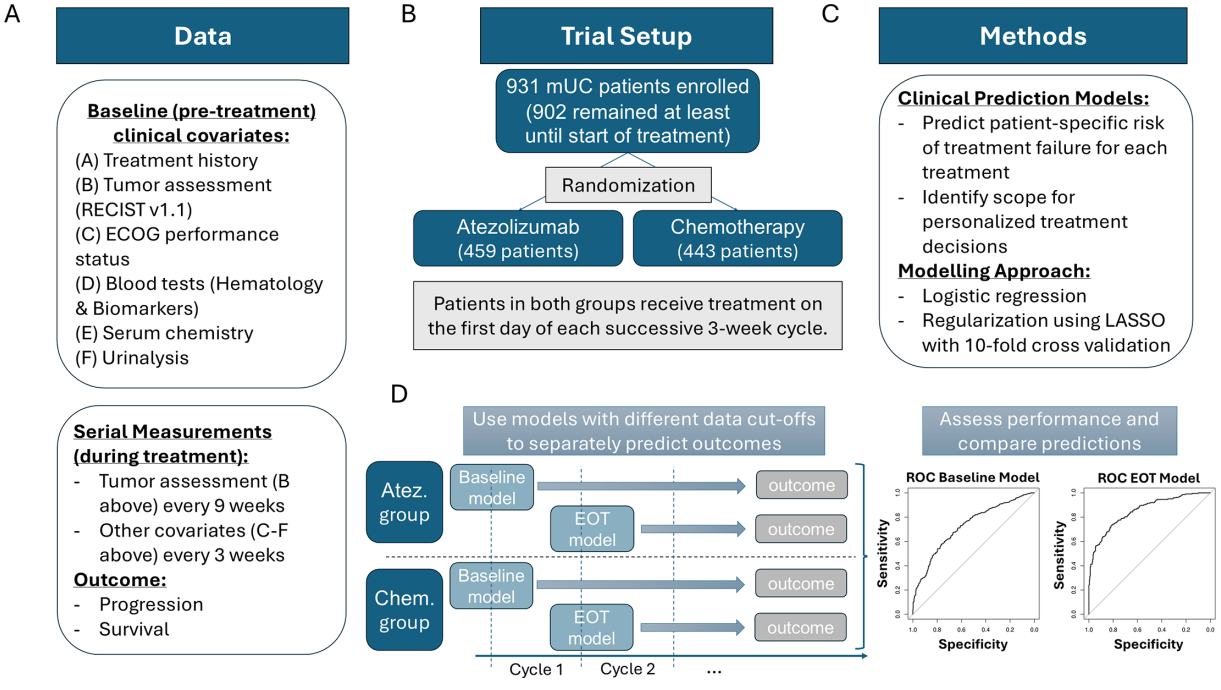

Figure 1 Schematic of our workflow. (A) Description of the available data from the phase three clinical trial IMVigor211. (B) Sketch of the trial composition of IMVigor211. (C) Brief summary of our methodology. (D) Sketch of our workflow. ECOG, Eastern Cooperative Oncology Group; LASSO, least absolute shrinkage and selection operator; mUC, metastatic urothelial carcinoma.

The IMvigor211 dataset contains information on patients' treatment history prior to the study. During the trial, tumor imaging and progression assessments based on RECIST V.1.1 were scheduled to occur in 9-week intervals (every three treatment cycles) up to week 54, then every 12 weeks. Other clinical covariates were collected in 3-week intervals, starting at baseline and aligning with the treatment cycles. These covariates included blood tests, comprehensive serum chemistry, and urinalysis (figure 1A, online supplemental tables 1 and 2). Additionally, adverse events were recorded as they occurred.

Statistical modeling approach

To predict the clinical endpoint of risk of treatment failure (defined in detail below) under either atezolizumab or chemotherapy, we established two logistic regression models using least absolute shrinkage and selection operator (LASSO) for regularization. The first model used only baseline variables, while the second model was based on both baseline and EOT variables (see "Methods" section). We used the area under the receiver operating characteristic curve (AUC) as a metric to assess the performance of these models relative to each other and relative to other known predictive factors.[11 12] Due to dropouts, we estimated the EOT model on only a subset of the patients in the baseline model. To disentangle changes in model performance stemming from added information versus patient selection, we additionally estimated a version of our baseline model in which we restricted the patient cohort to only those patients used in estimating the EOT model.

Our models predict treatment-specific risks for atezolizumab and chemotherapy for each patient. This approach allowed us to investigate the potential for personalized treatment decisions at baseline by comparing the mean predicted risk under each therapy to the mean predicted risk if every patient received the therapy that minimizes their personal predicted risk. Analogously, we compared predicted risks in the EOT model to quantify the benefits of switching between therapies after the first treatment cycle. Lastly, we characterized groups of patients who stood out as having very large differences in predicted risks between the two treatments.

Outcome variable and censoring considerations

The outcome of interest, treatment failure, is defined as having a best recorded treatment response of 'PD'. To account for dropouts within the first three cycles, we additionally classified patients as having treatment failure if they died prior to the first tumor assessment. Patients who dropped out of the study for other reasons before having a tumor assessment were, thus, not counted toward the treatment failure group. However, as a robustness check, we verified that including these patients in the treatment failure group yielded similar results for both models (online supplemental tables 3 and 4).

A second potential source of censoring bias is that patients who demonstrate early PD may actually have pseudoprogression with subsequent response, in which case they would be falsely counted toward the treatment failure group. We expect this effect to be relatively minor in advanced urothelial carcinoma, as pseudoprogression is a rare event.[15] Nevertheless, as a robustness check, we also estimated models in which we changed our outcome variable to "progression or death prior to a landmark time" (online supplemental tables 5 and 6).

Data preparation

Out of 931 patients, 29 patients dropped out before the start of treatment. The remaining 902 patients were included in our analysis regardless of when they discontinued treatment. To deal with missing values in the dataset, we excluded all cardinal variables with missing values at baseline. For missing categorical variables, we defined missing values as a separate category and verified, as a sensitivity analysis, that our prediction models did not depend on this category (online supplemental methods 1.1).

For the EOT model, we removed all variables that had more than 20 patients with missing values and excluded all patients from the estimation who had missing values in the remaining set of variables. This approach left us with 483 patients. The variables that were retained included some that were previously recognized as prognostic, for example, performance status, sites of metastasis, and hemoglobin, and others which are not commonly recognized as prognostic, for example, urine-specific gravity. The complete list of the variables in the baseline and EOT models is shown in online supplemental tables 1 and 2, respectively. Note that the added variables in the EOT model compared with the baseline model, with this exclusion procedure, include not only variables that were measured in the first treatment cycle, but also variables measured at baseline for which patients that were now excluded in the EOT model had missing values. Differences in performance between the models can, therefore, not be purely attributed to variables which are measured post baseline.

Even though treatment cycles in both treatment groups in IMVigor211 were scheduled to last 3 weeks, there was considerable variability between patients with respect to when clinical measurements for the first treatment cycle were taken. We, thus, chose day 22 to day 35 as the period for examining the EOT variables, to avoid having to drop patients whose measurements for the first treatment cycle were taken slightly later than was scheduled. Consequently, we only considered patients who survived beyond day 35 of the study in our EOT model.

Estimation procedure

All variables were centered at zero and scaled to have unit variance prior to estimation. For the logistic regression model, we incorporated each variable listed in online supplemental tables 1 and 2 independently, as well as in interaction with the treatment group variable. Thereby, we separated the effect of each covariate into a common

and a treatment-specific effect, which allowed us to identify covariates that differentially affect the risk of treatment failure in the two treatment groups. Moreover, for each variable in our EOT model (online supplemental table 2), we also included an interaction term with an indicator for whether a patient experienced any early immune-related adverse events (see discussion in online supplemental methods 1.2). Since our dataset contained hundreds of clinical variables, we used LASSO to select the most relevant covariates for our logistic regression models. To balance model complexity and model fit, we used 10-fold cross-validation to choose the LASSO penalty parameter.

RESULTS

Patient characteristics

At baseline, patients in the trial cohort had an average age of 66 years. All patients had previously received platinum-based chemotherapy (14.4% received both carboplatin and cisplatin, 28.5% only carboplatin, and 55.9% only cisplatin). Moreover, patients had an average of 1.95 metastatic sites, 28.5% had liver metastases, and 58.2% had lymph node metastases at baseline[13] (online supplemental table 1).

Those patient characteristics were very similar among the subset of patients we considered in our EOT model (see online supplemental table 2). The average age was 66, the percentages of patients having received prior treatment with carboplatin and cisplatin, carboplatin only, or cisplatin only were 15.7, 24.4, and 58.6 respectively, and the average number of metastatic sites was 1.92 with 24.2% having liver metastases and 60% having lymph metastases.

We also investigated whether there were any structural differences between the set of patients in our EOT model and the set of patients who lived beyond day 35 of the study, but who were excluded from the EOT model due to missing values. Based on the baseline variables without missing values, we found no such difference (F-test, p=0.107 (0.098) when performed on the first 30 (40) principal components of the data capturing 91% (99.8%) of the variability in the data).

Prognostic model using baseline parameters

We used a logistic regression modeling approach using baseline clinical and laboratory variables to predict the risk of progression as the best response after atezolizumab treatment in chemotherapy-refractory mUC (n=902 patients). This baseline model achieves an AUC of 0.725 across the two treatment groups, with a sensitivity of 0.39 at a specificity of 0.90 (figure 2A). When considering the chemotherapy group or the atezolizumab group separately, our models achieve AUCs of 0.696 and 0.717, respectively (figure 2C). Factors that were significantly associated with higher risk in both treatment groups include the presence of liver metastases (OR: 1.579, p<0.001), low PD-L1 expression (IC0/IC1; not IC2; OR: 1/0.834=1.199, p=0.015), low albumin levels (OR: 1.285, p=0.001), and low serum hemoglobin (OR: 1.210, p=0.015) (table 1). Additionally, we identified an association between the T-stage of the primary tumor and treatment outcomes, with a high T-stage (4a or 4b) being weakly but not significantly associated with worse outcomes for both therapies (ORs: 1.132 and 1.176, p=0.100 and 0.050), and a T1-stage predicting a lower risk in the atezolizumab group specifically (OR: 0.792, p=0.003).

Prognostic model using EOT parameters

We then estimated a logistic regression model that takes into account EOT variables. This EOT model achieved an AUC of 0.848 overall, with 0.871 for the atezolizumab treatment group and 0.788 for the chemotherapy treatment group (figure 2). At a specificity of 0.90 across both treatments, the sensitivity was 0.60. Compared with the baseline model, the EOT model had a much larger set of variables with non-zero coefficients in our LASSO approach, and no covariates stood out as dominant predictors (table 2). The set of selected variables includes various changes in values of laboratory tests over the first cycle of treatment; for instance, increases in erythrocytes are associated with a lower risk (OR: 0.681, p=0.031). Moreover, there is a large set of variables that are differentially associated with risk in both treatment arms (table 2). For instance, high urine-specific gravity and the presence of bone, but no other, metastases is associated with a higher risk of treatment failure for treatment with atezolizumab (OR: 1.504, p=0.009 and OR 1.458, p=0.036, respectively), while T-stage T1 is associated with a higher risk of chemotherapy treatment failure (OR: 0.761, p=0.035).

The increase in AUC of the EOT compared with the baseline model cannot be attributed solely to the added information that the additional EOT variables contribute, since the EOT model describes a selected subset of the population in the baseline models—those patients who survived beyond day 35 of the study. To isolate the magnitude of this improvement in AUC compared with the baseline model attributable to the additional covariates, we repeated the estimation procedure for the set of patients that were included in the EOT model, but with only the covariates that were available in our baseline model. This restricted model achieves an AUC of 0.760 (compared with the AUC of the baseline model of 0.725, and the set of variables selected through LASSO in this restricted model is relatively similar to the baseline model (see online supplemental table 7). The remaining sizeable gap in AUC between this restricted model and the EOT model (0.760 vs 0.848) suggests that the added covariates do provide substantial relevant further information. This observation is further supported by a notable reduction in the Akaike information criterion from 597.68 in the restricted model to 538.17 in the EOT model, indicating improved model fit while accounting for the increase in model complexity.

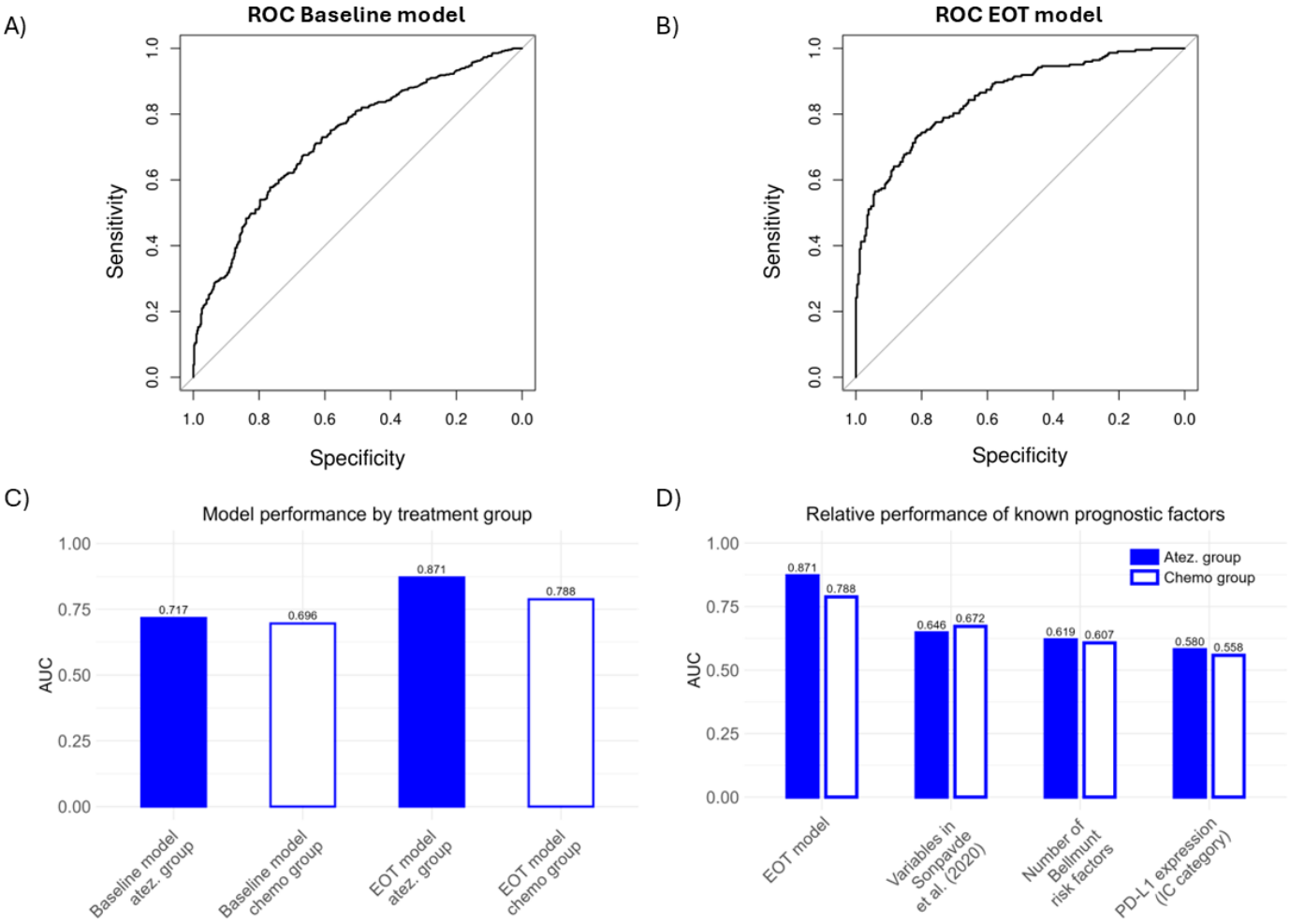

Figure 2 Model performance. (A, B) Receiver operating characteristic (ROC) curves of the baseline model and the EOT model for both treatment groups together. (C) Area under the ROC curves by treatment groups. (D) Area under the ROC curves of alternative logistic regression models estimated using the same patient cohort as for the EOT model, and using only known risk factors. From right to left, the models have the following regressors: categorical indicators for the PD-L1 expression categories 'IC0', 'IC1' and 'IC2'; the number of Bellmunt risk factors identified[12]; and the factors identified by Sonpavde *et al*,[11] including presence of liver metastases, ECOG performance status, platelet count, neutrophil-to-lymphocyte ratio and lactate dehydrogenase. AUC, area under the curve; EOT, early-on-treatment.

EOT model versus conventional baseline predictors

We then compared the prediction performance of our model to that of known baseline risk factors, namely those reported by Sonpavde *et al*,[11] the Bellmunt risk factors,[12] and PD-L1 expression, on the IMvigor211 dataset. To this end, we estimated three further logistic regression models with these known risk factors as regressors and without employing LASSO. Our EOT model outperforms each of these models by a large margin: while the models based on known risk factors achieve AUCs between 0.580 and 0.646, our EOT model achieves an AUC of 0.871 for the atezolizumab group. Similarly, for the chemotherapy group, the alternative models achieve AUCs between 0.558 and 0.672, while our EOT model achieves an AUC of 0.788 (figure 2D). Some of the components of the best-performing alternative model in figure 2D, namely presence of liver metastases as well as platelet and lymphocyte counts, also feature in our EOT model. Yet, the improved performance of our EOT model suggests that combining these components with further regressors substantially improves predictive ability.

At baseline, not all covariates in the Sonpavde *et al*[11] model are available for all patients. Models using the number of Bellmunt risk factors or PD-L1 expression (analogous to the other two alternative models in figure 2D) parameterized on the patient cohort in the baseline model yield AUCs of 0.612 and 0.540, respectively (not shown). Here, again, our baseline model outperforms these models (AUC of 0.725 across the two treatment groups).

We then further investigated the distributions of predicted risks of treatment failure in our models (figure 3). To illustrate the relationship between predicted risks and realized outcomes for the baseline and the EOT model, respectively, we visualized patient outcomes on a plane spanning the predicted risk of treatment failure of patients in the chemotherapy group (horizontal axis in figure 3A) and the predicted risk of treatment failure

Table 1 Baseline model

Variable	OR	95% CI (lower)	95% CI (upper)	P value	
Liver met.	1.579	1.355	1.840	<0.001	***
Albumin<=lower limit of normal	1.285	1.105	1.495	0.001	**
Atez. × TUM Stage T1	0.792	0.680	0.923	0.003	**
IC2 (5%–10% PD-L1 positive immune cells)	0.834	0.722	0.965	0.015	*
Serum hemoglobin <100 g/L	1.210	1.037	1.413	0.015	*
TUM stage T4B	1.176	0.999	1.384	0.050	
Atez. × Visceral met.	1.231	0.969	1.564	0.087	
TUM stage T4A	1.132	0.977	1.311	0.100	
Age	0.892	0.770	1.034	0.128	
Brain met.	0.887	0.760	1.035	0.128	
Atez. x Bone met. only	1.145	0.959	1.365	0.135	
Prior treatment with cisplatin and carboplatin	0.899	0.773	1.046	0.171	
Atez.	1.197	0.892	1.606	0.231	
Lymph met.	0.922	0.799	1.064	0.271	
Atez. × Prior treatment with only cisplatin	1.094	0.896	1.336	0.377	
Atez. x Distant metastasis stage M0	1.088	0.894	1.323	0.400	
(Intercept)	1.270	0.030	53.762	0.901	
Atez. × TUM stage TIS	2.832	4.644e-22	1.727e+22	0.968	

(*, **and *** indicate p values <0.5,0.01 and 0.001, respectively). Variables are ordered by p value. “H” and “L” indicate clinical annotations for high and low values, respectively.

of patients in the atezolizumab group (vertical axis in figure 3A). The position of an individual patient on this plane indicates the predicted risk of treatment failure under either treatment, while color and symbol, respectively, indicate the treatment that each patient received and their realized treatment outcome, respectively.

Comparison of the models in the atezolizumab versus chemotherapy groups

We found that the predicted risks of treatment failure for atezolizumab and chemotherapy are strongly correlated with each other in the baseline model (figure 3B, Pearson correlation coefficient of 0.86). This observation suggests a limited scope for improving outcomes with personalized treatment decisions. The predicted risks are less correlated in the EOT model (figure 3C, Pearson correlation coefficient of 0.77). Comparing the mean predicted risks of treatment failure for both treatments across all patients in our cohort, we found that at baseline, the risk of treatment failure can be reduced by 4.3 percentage points compared with the standard of care (atezolizumab) and by 3.4 percentage points compared with random assignment in the trial when switching to the alternative treatment. The potential for risk reduction suggested by our EOT model is larger: according to this model, personalized treatment decisions at day 35 reduce the risk of treatment failure by 6.6 percentage points relative to standard of care and by 6.2 percentage points relative to random assignment (figure 3C). Moreover, the EOT model suggests that 33.7 percent of patients who were initially treated with atezolizumab would benefit from switching to chemotherapy after the first treatment cycle and predicts that such a switch reduces the risk of treatment failure for the subpopulation of patients who survive at least 35 days by 15.6 percentage points (from 67.1% to 51.5%).

Based on this analysis, we observed that there are several patients with substantially (more than 30 percentage points) higher predicted risk under atezolizumab than under chemotherapy (figure 3C). This observation agrees with observed treatment outcomes—most patients in this region who received atezolizumab had treatment failure, while those who received chemotherapy did not. Conversely, there are far fewer patients with a more than 30 percentage points higher predicted risk under chemotherapy than under atezolizumab (figure 3C). The patients with the largest differences in predicted risk of treatment failure between the two treatments, therefore, tend to have a higher predicted risk under atezolizumab and a lower risk under chemotherapy, and not vice versa. This finding is consistent with clinical observations[1] that there is more variability in responses to immunotherapies compared with chemotherapies, and that this higher variability is driven by a combination of (1) moderately improved outcomes among responders and (2) a group of patients who show no response.

Differentially predictive covariates for response to atezolizumab versus chemotherapy

In the EOT model, differences in predicted risks between the two treatments were associated with covariates with treatment-specific effects (table 2). We, thus, investigated which of these covariates are particularly enriched among

Table 2 EOT model

Variable	OR	95% CI (lower)	95% CI (upper)	P value	
Atez. × urine specific gravity	1.504	1.110	2.038	0.009	**
Liver met.	1.595	1.121	2.270	0.009	**
Protein	0.728	0.562	0.943	0.016	*
Erythrocytes	0.724	0.551	0.951	0.020	*
ΔErythrocytes	0.681	0.481	0.965	0.031	*
Atez. × TUM stage T1	0.761	0.591	0.980	0.035	*
Calcium “H”	1.343	1.021	1.767	0.035	*
Atez. × Bone met. only	1.458	1.024	2.075	0.036	*
Platelets	1.377	1.020	1.859	0.036	*
TUM stage T4B	1.334	1.008	1.765	0.044	*
Age	0.772	0.596	1.000	0.049	*
Thyroxine (free)	0.682	0.465	0.999	0.050	
Prior treatment with cisplatin and carboplatin	0.770	0.571	1.037	0.084	
IC2 (5%–10% PD-L1 positive immune cells)	0.805	0.628	1.032	0.087	
Calcium “L”	1.336	0.948	1.883	0.097	
Lymphocytes “L”	1.320	0.941	1.854	0.108	
Prior treatment with only cisplatin	1.273	0.948	1.707	0.109	
Thyrotropin	0.825	0.649	1.050	0.118	
ΔMagnesium	1.487	0.874	2.530	0.143	
Transitional cell carcinoma with mixed histology	1.185	0.931	1.508	0.166	
EIRAE × ΔMagnesium	1.241	0.898	1.715	0.190	
Magnesium “H”	1.242	0.873	1.768	0.226	
Atez. × Bilirubin	1.339	0.797	2.251	0.269	
Atez. × Aspartate aminotransferase	1.298	0.791	2.132	0.301	
EIRAE	1.177	0.830	1.668	0.359	
Atez × glucose	1.196	0.788	1.816	0.401	
Atez. × Liver met.	1.189	0.792	1.784	0.403	
Visceral met.	1.158	0.820	1.636	0.403	
Leucocytes	1.092	0.843	1.414	0.506	
Atez. × ΔBilirubin	1.155	0.755	1.767	0.508	
ΔAlkaline phosphatase	1.103	0.784	1.551	0.571	
Prothrombin international normalized ratio	1.155	0.694	1.922	0.578	
Prothrombin international normalized ratio “H”	1.125	0.683	1.855	0.644	
Atez.×Lymphocytes “L”	1.093	0.747	1.599	0.645	
ΔC reactive protein	0.728	0.180	2.943	0.656	
Atez. × Visceral met.	0.942	0.554	1.602	0.824	
Chloride	0.986	0.766	1.270	0.915	
Atez. × Δ25-dihydroxyvitamin D	0.049	1.258e−216	1.920e+213	0.990	
Atez. × TUM stage TIS	5.124	1.852e−119	1.418e+120	0.991	
Atez. × ΔThyrotropin	3.838	1.656e−165	8.897e+165	0.994	
Sodium “H”	2.298	1.355e−153	3.896e+153	0.996	
(Intercept)	0.911	4.098e−23	2.026e+22	0.997	

(*, **and *** indicate p values below 0.05, 0.01 and 0.001). Variables are ordered by p value. “H” and “L” indicate clinical annotations for high and low values, respectively.
EIRAE, early immune-related adverse event; EOT, early-on-treatment.

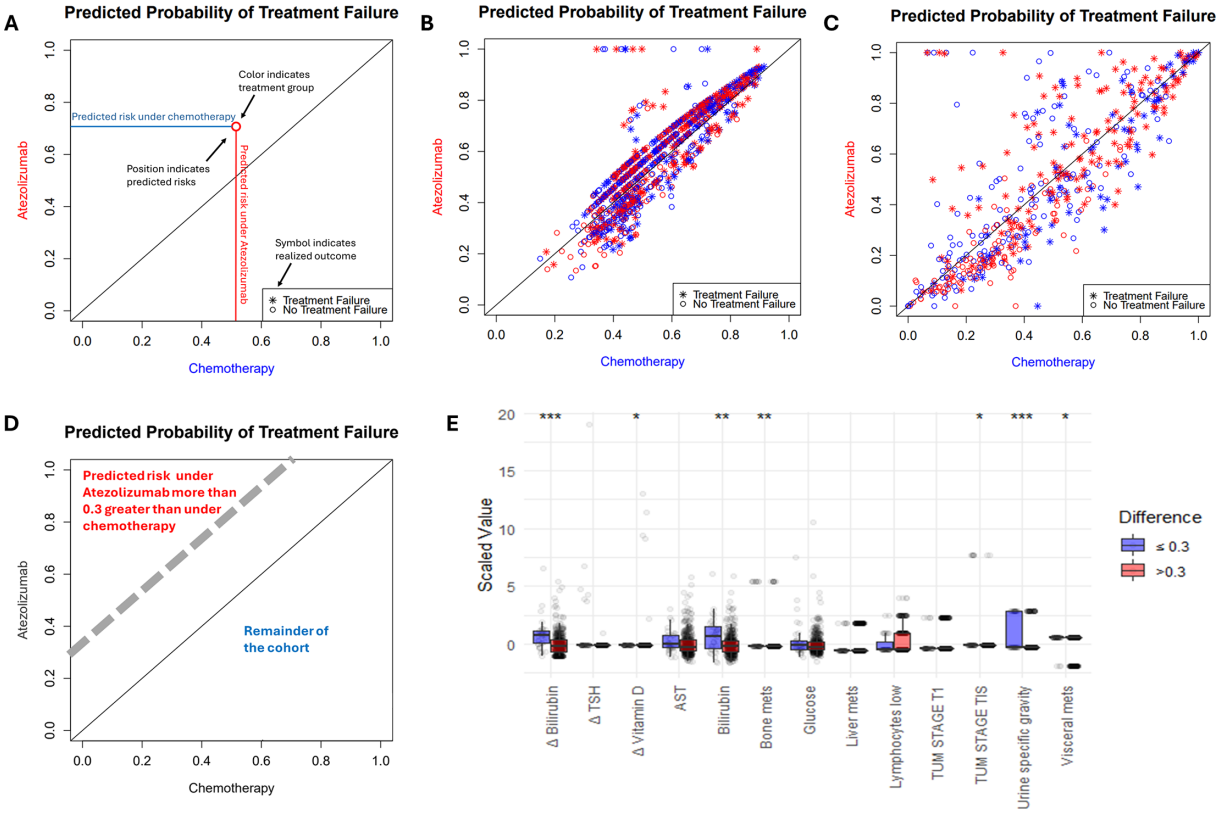

Figure 3 Model predictions. (A) Conceptual guide to the plots in B, C. (B, C) Scatter plots of predicted risk of treatment failure for all 902 patients in the baseline model (B) and all 483 patients in the EOT model (C). (D) Schematic demonstrating how the patient cohort was split for analyses shown in E. (E) Distribution of the variables in the EOT model which have a treatment-specific effect, for patients whose risk of treatment failure is at least 30 percentage points higher under atezolizumab than under chemotherapy (red) and for the remaining patients (blue). AST, aspartate aminotransferase; EOT, early-on-treatment; TSH, Thyroid-Stimulating Hormone .

the patients with the largest differences in predicted risks of treatment failure due to the two treatment options to characterize this set of patients. We dichotomized the patient cohort into two groups based on whether the patients' risk of treatment failure under atezolizumab was more than 30 percentage points greater than under chemotherapy (figure 3D). We then contrasted the respective distributions of the covariates with treatment-specific effects from our EOT model for those two groups (figure 3E). We found that there is no single covariate that clearly delineates these two groups—rather, there are several covariates whose means differ significantly between the groups, but the distributions of those covariates strongly overlap. This characterization suggests that a multidimensional score is needed to characterize those patients who would fare better with chemotherapy than with atezolizumab.

Analogously, we investigated the group of patients for whom our EOT model predicts a more than 20 percentage points higher risk of treatment failure under chemotherapy (online supplemental figure 1). We found significant differential enrichment for all except three of the covariates (changes in vitamin D, glucose levels, and the indicator for low levels of lymphocytes) that have a treatment-specific effect in the EOT model, showing that also for this group of patients the difference in predicted outcomes between the two treatments is driven by several different covariates. For instance, aspartate aminotransferase levels, bilirubin levels, bilirubin changes over the first cycle, and urine-specific gravity are on average significantly higher among patients with more than 20 percentage points higher risk of treatment failure under chemotherapy.

DISCUSSION

While immune checkpoint inhibitors can yield very favorable clinical responses in some patients, there is also a substantial share of patients who show no positive clinical response. This large variability in treatment outcomes has sparked various efforts to identify predictive markers to distinguish patients who will respond to immunotherapy from those who will not,[10 11] with the goal of identifying patients with a high risk of early treatment failure on immunotherapy who might have better expected treatment outcomes by switching to other therapeutic options.

We set out to directly quantify the opportunity for improved treatment outcomes by simultaneously estimating a patient-specific risk of treatment failure for atezolizumab and chemotherapy. We analyzed readily

available baseline and EOT clinical and laboratory factors from the IMvigor211 phase III trial which compared atezolizumab versus chemotherapy for patients with mUC that progressed following platinum-based chemotherapy.

Our baseline model achieved an AUC of 0.725 overall, 0.717 for the atezolizumab group and 0.696 for the chemotherapy group, while the EOT model achieved an AUC of 0.848 overall, 0.871 for the atezolizumab group and 0.7879 for the chemotherapy group. Thus, the EOT model substantially outperforms the baseline model and achieves a substantial improvement in prediction performance relative to currently available prediction tools, as the comparisons with models based on known predictors demonstrate. Further improvements in predictive performance might be achievable when training models on larger and multimodal datasets of relevant patient groups.

While the covariates in our models are derived from routine clinical measurements available in many countries, access to immune checkpoint inhibitors remains limited in some regions. By improving the precision of patient selection, we hope that our model will not only improve treatment outcomes but also contribute to making immune checkpoint inhibitors more cost-effective and accessible.

The EOT model included some factors that were present in the baseline model, for example, age and indicators for T-Stage T4B and for the presence of liver metastases, but most factors were newly identified clinical and laboratory factors, such as urine specific gravity, protein and erythrocyte levels, as well as the change in erythrocyte levels over the first cycle. Several factors, such as urine-specific gravity and the indicator for T-Stage T1, differentially predict treatment in atezolizumab versus chemotherapy. It is intriguing that we observed several patients with predicted high risk of early progression on atezolizumab and low risk of progression on chemotherapy, while the reverse was not observed.

For many factors in our model, an association with treatment outcomes is biologically plausible; for instance, we indeed expect that the presence of metastases in different organs correlates with worse treatment outcomes. For other factors, the biological link to treatment outcomes is less apparent; for instance, the T-stage of the primary tumor features prominently in our model, even though the model was built on a cohort of patients with metastatic disease. While our findings can be used to formulate hypotheses for mechanistic links between factors and outcomes, for example, a higher T-stage may be associated with higher rates of ongoing seeding of new metastases attributable to greater molecular heterogeneity and evolution of a larger primary tumor, the associations in our model are all correlative and not necessarily causal.

One prominent predictor in both the baseline and the EOT model is PD-L1 expression. While this feature is mechanistically linked to atezolizumab response, PD-L1 expression is known to fluctuate dynamically, often limiting its predictive utility.[16 17] The performance of predictive models would, therefore, likely improve with more granular longitudinal PD-L1 measurements.

Our findings based on the IMVigor211 patient cohort indicate that the risks of treatment failure under atezolizumab and chemotherapy are strongly correlated at baseline and less correlated in the EOT model (Pearson correlation coefficients of 0.84 and 0.77, respectively). Accordingly, the EOT model predicts larger differences in risk of treatment failure between the two treatments, and thus, more scope for personalized treatment decisions to improve treatment outcomes. In particular, the EOT model suggests that switching to chemotherapy reduces the risk of treatment failure for about a third of patients treated with atezolizumab. The resulting average reduction in risk for this group (67.1%–51.5%) is appreciable but not substantial. However, the predicted reduction in risk within this group has a wide distribution, with large reductions (>30 percentage points) for many patients. Identifying such patients may substantially improve treatment outcomes for this subset.

These suggestions require validation in a better controlled setting, as the EOT model is parameterized with a share of the observations gathered at a time when individuals in the two treatment arms have undergone one cycle of treatment, so that the two arms are no longer perfect counterfactuals of one another. Notably, circulating tumor (ct)-DNA appears promising to detect the molecular disease burden to trigger adjuvant therapy and early response to therapy. However, it is unclear if decisions regarding changes in ongoing therapy for metastatic disease can be made based on molecular disease burden progression.[18] Moreover, the requirement for adequate amounts of tumor material, turnaround time, and costs are barriers to implementing tumor-informed ctDNA assays. One ongoing phase III trial, SERENA-6 (NCT04964934), is evaluating the efficacy of switching from an aromatase inhibitor to camizestrant, an oral selective estrogen receptor degrader, while continuing the same CDK4/6 in combination, if molecular progression is observed in terms of a new *ESR1* mutation detected in ctDNA for HR+/HER2– advanced breast cancer.[19] However, this trial does not address other mechanisms of progression. Another previously reported clinical trial evaluated a switch from initial chemotherapy to a different chemotherapy based on early changes in circulating tumor cells (CTCs) in patients with metastatic breast cancer.[20] While this previously reported trial did not identify improved outcomes by employing information for early CTC dynamics, a switch to a different class of therapy, for example, immunotherapy or targeted therapy, may have been more successful.

We found that our ability to predict treatment failure increases substantially when moving from the baseline to the EOT model (AUC increase from 0.725 to 0.848). Aside from the added information contained in the additional variables in the EOT model, part of this increase is due to differences in the composition of the patient cohorts at baseline and at EOT. A model using baseline covariates

that is estimated using only the subset of patients who are considered in the EOT model performs better than the baseline model with the full patient cohort (AUC increase from 0.725 to 0.76).

Our modeling of early dynamics under treatment was constrained by the time intervals between clinical measurements in IMVigor211. Our EOT model consequently predicts refractory disease after 1 cycle of treatment, which is roughly 1–2 months before refractory disease would likely be diagnosed via imaging. Nevertheless, identifying primary refractory disease before clinical or radiographic progression has value since clinical progression may lead to a decline in performance status and ineligibility for switching to an alternative systemic therapy. While reducing the duration of futile treatment by this margin is valuable, we expect that the power to predict treatment failure based on the variables in the EOT model increases gradually from the beginning of treatment, so that shorter time intervals between measurements might enable even earlier identification of patients likely to have treatment failure.

Our study is a retrospective analysis, although prospectively collected data from a well-conducted phase III trial were used. A key difference between the patient populations in the baseline compared with the EOT model is that at baseline, there are far more patients who will die before ever having their treatment response status assessed. As these patients are counted toward the treatment failure group in the baseline model, the increased prediction performance of baseline covariates for the patient cohort in the EOT model relative to the patient cohort in the baseline model indicates that there is limited overlap between factors that predict survival beyond 35 days and those that predict treatment response.

Although atezolizumab has been withdrawn from the urothelial carcinoma indication, leveraging this trial data offers a unique opportunity to provide proof of principle by studying treatment responses to a PD-L1 inhibitor and the alternative chemotherapy in the same cohort. Given the similarities of PD1/L1 inhibitors, our modeling approach is easily transferrable to other PD-L1 inhibitors and potentially to other classes of agents.

Our approach differs from previous prediction models for treatment outcomes under atezolizumab in that we do not preselect the number of variables that are used for prediction. To avoid overfitting, we use LASSO for regularization and choose the LASSO parameter using 10-fold cross-validation. With this approach, our model combines a large set of variables, and no individual variables stand out as dominant predictors. Similarly, the group of patients who have a substantially lower risk of treatment failure under chemotherapy than under atezolizumab cannot easily be delineated by a single variable, as several variables contribute to this difference in risk.

The performance of our EOT model (AUC=0.848) suggests that a large multivariable model using routine EOT clinical and laboratory metrics has the potential to improve our ability to predict treatment outcomes for patients receiving atezolizumab. However, validation on independent patient cohorts and evaluation in other therapeutic contexts is needed.

Our approach provides a proof of principle that dynamics early during treatment can be used to simultaneously predict the risk of refractory disease for PD-L1 inhibitors and chemotherapy, and thus enable personalized therapeutic decisions.

For the IMVigor211 trial dataset, our EOT model suggests that ~34% of patients who were initially treated with atezolizumab may benefit from switching to another therapy after the first treatment cycle, which may reduce the risk of early treatment failure, decline in performance status, and ineligibility for subsequent treatment. This early setting after one cycle of therapy may represent an opportunity to develop novel systemic therapy employing early dynamics of routine clinical and laboratory parameters, which is considerably more cost-effective and globally scalable compared with molecular assays such as ctDNA.

Author affiliations
[1]Department of Data Science, Dana-Farber Cancer Institute, Boston, Massachusetts, USA
[2]Department of Biostatistics, Harvard T H Chan School of Public Health, Cambridge, Massachusetts, USA
[3]Department of Stem Cell and Regenerative Biology, Harvard University, Boston, Massachusetts, USA
[4]Center for Cancer Evolution, Dana-Farber Cancer Institute, Boston, Massachusetts, USA
[5]University of Central Florida, Orlando, Florida, USA
[6]AdventHealth Cancer Institute, Orlando, Florida, USA
[7]The Eli and Edythe L Broad Institute, Cambridge, Massachusetts, USA
[8]Ludwig Center at Harvard, Harvard Medical School, Boston, Massachusetts, USA

X Guru Sonpavde @sonpavde

Acknowledgements This publication is based on research using data from data contributors Roche that has been made available through Vivli, Inc. Vivli has not contributed to or approved, and is not in any way responsible for, the contents of this publication. While institutional affiliations are limited to two per journal policy, Franziska Michor and Thomas O McDonald would like to acknowledge their affiliation with the Center for Cancer Evolution at the Dana-Farber Cancer Institute, and Franziska Michor would additionally like to acknowledge her affiliation with the The Eli and Edythe L. Broad Institute and with the Ludwig Center at Harvard.

Contributors CG, GS and FM conceptualized and designed the study. GS facilitated and funded data access. CG conducted the analyses in discussion with all other authors. CG and FM wrote the manuscript with edits from GS, TOM and PJC. Guarantor: CG.

Funding This study was supported by the Center for Cancer Evolution at the Dana-Farber Cancer Institute.

Competing interests GS declares the following conflicts of interest. Advisory Board: EMD Serono, BMS, Merck, Seattle Genetics, Astellas, Janssen, Bicycle Therapeutics, Pfizer, Gilead, Scholar Rock, Eli Lilly, Loxo Oncology, Vial, Aktis, Daiichi-Sankyo; Consultant/Scientific Advisory Board (SAB)/trial steering committee: Syapse, Merck, Servier, Syncorp, Ellipses; Research Support to institution: EMD Serono, Jazz Therapeutics, Bayer, Sumitomo Pharma, Blue Earth Diagnostics; Speaker: Seagen, Gilead, Natera, Exelixis, Janssen, Astellas, Bayer, Aveo, Pfizer, Merck; Data safety monitoring committee (honorarium): Mereo; Employment: Spouse employed by Myriad, Exact Sciences; Travel: BMS, Astellas. FM is a co-founder of and has equity in Harbinger Health, has equity in Zephyr AI, and serves as a consultant for both companies. She is also on the board of directors of Recursion Pharmaceuticals. She declares that none of these relationships are directly or indirectly related to the content of this manuscript. The other authors declare no conflicts of interest.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available.

ORCID iDs
Christopher J Graser http://orcid.org/0009-0003-8552-7767
Thomas O McDonald http://orcid.org/0000-0002-5272-0187
Paul J Catalano http://orcid.org/0000-0002-1105-8852
Guru Sonpavde http://orcid.org/0000-0002-1010-9611
Franziska Michor http://orcid.org/0000-0003-4869-8842

REFERENCES

1 Bellmunt J, de Wit R, Vaughn DJ, *et al*. Pembrolizumab as Second-Line Therapy for Advanced Urothelial Carcinoma. *N Engl J Med* 2017;376:1015–26.
2 Patel MR, Ellerton J, Infante JR, *et al*. Avelumab in metastatic urothelial carcinoma after platinum failure (JAVELIN Solid Tumor): pooled results from two expansion cohorts of an open-label, phase 1 trial. *Lancet Oncol* 2018;19:51–64.
3 Powles T, Eder JP, Fine GD, *et al*. MPDL3280A (anti-PD-L1) treatment leads to clinical activity in metastatic bladder cancer. *Nature New Biol* 2014;515:558–62.
4 Gibney GT, Weiner LM, Atkins MB. Predictive biomarkers for checkpoint inhibitor-based immunotherapy. *Lancet Oncol* 2016;17:e542–51.
5 Taube JM, Klein A, Brahmer JR, *et al*. Association of PD-1, PD-1 ligands, and other features of the tumor immune microenvironment with response to anti-PD-1 therapy. *Clin Cancer Res* 2014;20:5064–74.
6 Le DT, Durham JN, Smith KN, *et al*. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* 2017;357:409–13.
7 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. *The Lancet* 2016;387:1909–20.
8 Teo MY, Seier K, Ostrovnaya I, *et al*. Alterations in DNA Damage Response and Repair Genes as Potential Marker of Clinical Benefit From PD-1/PD-L1 Blockade in Advanced Urothelial Cancers. *J Clin Oncol* 2018;36:1685–94.
9 Pond GR, Niegisch G, Rosenberg JE, *et al*. New 6-factor prognostic model for patients (pts) with advanced urothelial carcinoma (UC) receiving post-platinum atezolizumab. *JCO* 2018;36:413.
10 Nassar AH, Mouw KW, Jegede O, *et al*. A model combining clinical and genomic factors to predict response to PD-1/PD-L1 blockade in advanced urothelial carcinoma. *Br J Cancer* 2020;122:555–63.
11 Sonpavde G, Manitz J, Gao C, *et al*. Five-Factor Prognostic Model for Survival of Post-Platinum Patients with Metastatic Urothelial Carcinoma Receiving PD-L1 Inhibitors. *J Urol* 2020;204:1173–9.
12 Bellmunt J, Choueiri TK, Fougeray R, *et al*. Prognostic factors in patients with advanced transitional cell carcinoma of the urothelial tract experiencing treatment failure with platinum-containing regimens. *J Clin Oncol* 2010;28:1850–5.
13 Powles T, Durán I, van der Heijden MS, *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. *The Lancet* 2018;391:748–57.
14 Study details | a study of atezolizumab compared with chemotherapy in participants with locally advanced or metastatic urothelial bladder cancer [IMvigor211]. ClinicalTrials.gov. Available: https://clinicaltrials.gov/study/NCT02302807 [Accessed 6 Oct 2024].
15 Soria F, Beleni AI, D'Andrea D, *et al*. Pseudoprogression and hyperprogression during immune checkpoint inhibitor therapy for urothelial and kidney cancer. *World J Urol* 2018;36:1703–9.
16 Davis AA, Patel VG. The role of PD-L1 expression as a predictive biomarker: An analysis of all US food and drug administration (FDA) approvals of immune checkpoint inhibitors. *J Immunother Cancer* 2019;7:1–8.
17 Grossman JE, Vasudevan D, Joyce CE, *et al*. Is PD-L1 a consistent biomarker for anti-PD-1 therapy? The model of balstilimab in a virally-driven tumor. *Oncogene* 2021;40:1393–5.
18 Powles T, Assaf ZJ, Davarpanah N, *et al*. ctDNA guiding adjuvant immunotherapy in urothelial carcinoma. *Nature New Biol* 2021;595:432–7.
19 Turner N, Huang-Bartlett C, Kalinsky K, *et al*. Design of SERENA-6, a phase III switching trial of camizestrant in *ESR1*-mutant breast cancer during first-line treatment. *Future Oncol* 2023;19:559–73.
20 Smerage JB, Barlow WE, Hortobagyi GN, *et al*. Circulating tumor cells and response to chemotherapy in metastatic breast cancer: SWOG S0500. *J Clin Oncol* 2014;32:3483–9.