Monitoring efficacy of checkpoint inhibitor therapy in patients with non-small-cell lung cancer

Annett Schiwitza¹, Hans-Ulrich Schildhaus², Birgit Zwerger³, Josef Rüschoff⁴, Christian Reinhardt³, Andreas Leha⁵, Stefan Andreas^{1,3} & Achim Rittmeyer^{*,3}

¹Department of Pneumology (Research & Teaching), Universitätsmedizin Göttingen, Göttingen, Germany

²Department of Pathology, Universitätsmedizin Göttingen, Göttingen, Germany

³LKI Lungenfachklinik Immenhausen, Immenhausen, Germany

⁴Department of Pathology, Pathologie Nordhessen, Kassel, Germany

⁵Department of Medical Statistics, Universitätsmedizin Göttingen, Göttingen, Germany

*Author for correspondence: Tel.: +49 5673 501421; Fax: +49 5673 501319; arittmeyer@lungenfachklinik-immenhausen.de

Aim: Radiological criteria alone do not reflect the entire population benefitting from checkpoint inhibitor therapy (CIT). This study aimed to detect patterns to assess CIT efficacy in non-small-cell lung cancer (NSCLC) patients. **Materials & methods:** We evaluated clinical, radiological and laboratory parameters in a retrospective cohort of NSCLC patients treated with nivolumab. **Results:** A total of 51 patients were included in the analysis. Most single parameters failed to reflect treatment benefit. Three laboratory parameters (lactate dehydrogenase, C-reactive protein and the neutrophil/lymphocyte ratio) combined in a weighted score could predict benefit with a sensitivity of 92.3% and a hazard ratio of 0.31 (95% CI: 0.16–0.59) in an early phase of therapy. Sorting patients by score showed a 1-year survival of 36% in those predicted as not benefitting versus 68% in those predicted to benefit. **Conclusion:** A weighted score integrating common serum markers could help detect patients benefitting from checkpoint inhibitors during ongoing CIT.

First draft submitted: 24 February 2019; Accepted for publication: 1 May 2019; Published online: 22 May 2019

Keywords: checkpoint inhibition • immunotherapy • monitoring efficacy • nivolumab • NSCLC • PD-1/PD-L1 inhibition

Lung cancer is the leading cause of cancer death worldwide, with about 1.8 million deaths every year [1,2]. Tyrosine kinase inhibitors can be employed in patients with a targetable driver mutation; however, as these mutations are rare, first-line chemotherapy with a platinum doublet remains the standard of care for most lung cancer patients. As pemetrexed is already used in first-line combinations in patients with nonsquamous non-small-cell lung cancer (NSCLC) and not approved in squamous cell lung cancer, chemotherapy with docetaxel is the standard of care for most patients receiving second-line treatment [3–6]. Only limited advances have been made over recent years for patients with no detectable driver mutations [7,8].

Checkpoint inhibition with drugs that inhibit the interaction of the programmed death ligand PD-L1 receptor on cytotoxic lymphocytes and PD-L1 as the corresponding ligand on tumor cells and many immune cells has led to an unprecedented improvement of overall survival (OS) in NSCLC and has become a new paradigm in the second-line treatment of NSCLC [9–12]. Phase III trials comparing nivolumab, pembrolizumab and atezolizumab with docetaxel in previously treated NSCLC revealed a median OS improvement of 3–4 months and a 12–22% increase in 12-month survival rates [4,5,13,14]. The safety profile of these checkpoint inhibitors was found to be favorable compared with standard chemotherapy. These results have already led to the approval of all three drugs.

Atypical and delayed responses can occur during treatment with immune checkpoint inhibitors, possibly due to invasion of an immune infiltrate increasing the tumor volume temporarily, but eventually leading to the death of cancer cells and delayed tumor shrinkage [15–17]. Immune-RECIST (iRECIST) criteria lead to about 1–2% higher

Future . Medicine







Figure 1. CONSORT flow diagram of the study. A total of 51 patients, who received study treatment consisting of at least four cycles of nivolumab and the corresponding CT scans, were included in the final analysis.

response rates than classical RECIST version 1.1 criteria, but neither RECIST nor iRECIST reflects the complete benefit achieved by checkpoint inhibitor therapy (CIT) [4,5,18,19].

The aim of this study was to identify differences in those patients that showed progressive disease at the first CT scan (or at least no real response) and try to find patterns that could differentiate pseudoprogressive disease from real progression. We retrospectively collected data from patients treated with nivolumab at our institution in order to evaluate clinical, pathological, radiological and laboratory results to identify repetitive signals that could help to assess the real benefit derived from therapy with checkpoint inhibitors.

Materials & methods

Participants

All consecutive patients (n = 68) with an Eastern Cooperative Oncology Group performance status of 0 or 1 treated between April 2015 and September 2016 with nivolumab off-protocol as second- or further-line treatment after failure of platinum-based chemotherapy at LKI Lungenfachklinik Immenhausen (Immenhausen, Germany) were included in this study. Those patients (n = 51), who had at least one CT scan after the first four cycles of nivolumab, were eligible for inclusion in the retrospective analysis (Figure 1).

Every participant provided signed written informed consent at the time of first lung cancer diagnosis to permit collection and scientific analyses of clinical, anonymized data. The trial was approved by the local ethics committee (Ethikkommission der Universitätsmedizin Göttingen; approval DOK_82_2016).

Procedure

Nivolumab was given as approved at 3 mg/kg by intravenous injection every 2 weeks. The treatment was continued for as long as the patient derived benefit from nivolumab at the discretion of the treating physician or until intolerable side-effects occurred.

Analyses

Blood samples (cell count, C-reactive protein [CRP], lactate dehydrogenase [LDH], potassium, sodium, calcium, urea, creatinine, aspartate transferase, alanine transferase, γ -glutamyl transpeptidase, alkaline phosphatase, bilirubin, total protein, glucose, uric acid, thyroid stimulating hormone and coagulation) and clinical symptoms (e.g., myalgia, fever, dyspnea, cough, weight change and pain) were analyzed and registered according to routine clinical procedures. We also extracted histological and cytological features, PD-L1 expression, and driver mutations if available.

PD-L1 biomarker analysis

PD-L1 protein expression was evaluated in pretreatment biopsies with an immunohistochemistry assay employing the DAKO clone 28-8 (Agilent, Waldbronn Germany). The stains were independently reviewed by two authors (H-U Schildhaus and J Rüschoff). PD-L1 results were given as the percentage of tumor cells (tumor proportion score) and the area of the immune infiltrate expressing PD-L1 independently of the degree of expression.

Table 1. Scoring scheme according to the measured biomarkers.							
Parameter	Score						
	2	1	0				
Relative CRP value change, % (×0.5) †	<-60	-60 to -20	≥-20				
Relative LDH value change, % (×2.0) †	<-10	-10 to 0	≥ 0				
Relative NLR value change, %	<-30	-30 to 10	≥10				
Relative tumor size change, %	<-10	-10 to 0	≥ 0				
Baseline value CRP, mg/l	<5	5 to 60.4	≥60.4				
Baseline value LDH, U/I	<150	150 to 207	≥207				
Baseline tumor size, mm	<32	32 to 61	≥61				

The final score was calculated by summation of the subscores for each parameter.

[†]The relative CRP value change score was multiplied by 0.5 and the relative LDH value change score was multiplied by 2.0. Therefore, the maximum score was 14.5 and the minimum score was 0.0.

Response assessment

Tumor response was evaluated according to RECIST version 1.1 and iRECIST [18,19]. Pseudoregression (PsPr) was defined as an increase of more than 20% in tumor size at the first CT scan and an objective remission (reduction of 30% or greater) after eight cycles of nivolumab, corresponding to iRECIST unconfirmed progressive disease (iUPD) at the first control and iRECIST partial response at the second. Objective response (OR) was defined as partial remission according to RECIST version 1.1, corresponding to a partial or complete response (iRECIST partial response or iCR) according to iRECIST. Stable disease (SD) was defined as described by RECIST version 1.1, corresponding to iso according to iRECIST. Real progressive disease (RPD) was defined as a continuous increase in tumor size of at least 20% after eight cycles of nivolumab, corresponding to iRECIST iUPD at the first CT scan and iRECIST confirmed progression at the second. We used RPD to distinguish progressive disease in those with pseudoprogressive disease (i.e., progression followed by remission) from (real) progressive disease (i.e., progression).

The CT scans were performed after every fourth cycle of nivolumab between day 43 and 56 and day 99 and 113 (Appendix Figure A.1). The corresponding laboratory values and clinical parameters were recorded on days 43 and 99 (\pm 3 days); accordingly, the CT scans were also referred to as day 43 and 99. If the second CT scan (day 99) showed progressive disease, this was regarded as RPD and the patient was discontinued from nivolumab. If a patient died after the first CT scan but before the second CT scan could be performed, this was also regarded as RPD.

Score development

We attempted to combine various parameters analyzed in order to produce a more robust and earlier signal from the laboratory values that showed some correlation with response (see results). Therefore, a predictor score was computed by integrating four parameters (LDH, the neutrophil/lymphocyte ratio [NLR], CRP and tumor size) at baseline as well as day 43, in other words, after four cycles of nivolumab and at the time of the first CT control. A score was allocated to each individual parameter based on the laboratory values determined. The relevant scores for the individual parameters were summed to give the final total score. Details of the chosen parameters and the methodology used to calculate the score are given in Table 1. The maximum total score was 14.5 and the minimum total score was 0.0 (note that a weighting was applied whereby the relative CRP value change score was multiplied by 0.5 and the relative LDH value change score was multiplied by 2.0).

As a first step in the development, we trained data from days 1 (baseline), 57, 71, 85 and 99 of CRP, LDH, the NLR, and change in sum diameter of target lesions assessed by CT scan. The performance of this predictor was assessed by means of a stratified tenfold cross-validation repeated ten-times and by receiver operating characteristic (ROC) curve analysis (see results). We then evaluated the score using data from baseline and day 43 in order to predict benefit from CIT as early as possible. The accuracy, area under the ROC curve [AUC] and sensitivity at a specificity of 50% were reported. Total 50% specificity was chosen in order to detect as many patients as possible benefitting from CIT. False positive assessments by the score were deemed of less importance in that context.

Statistics

The clinical variables for each of the response groups were summarized separately by mean \pm SD, continuously scaled, or absolute and relative frequencies, nominally scaled. Each clinical variable was compared between the response groups by analysis of variance, Fisher's exact test, or the likelihood ratio test. The relative change versus baseline was calculated for all potential biomarkers at each time point. Pairwise Mann–Whitney U-tests were used to compare the scores for day 43 and 99 between the RPD group and the other groups. Holm adjusted p-values from pairwise comparisons against RPD tests were determined. The OS was analyzed via two separate Cox regressions: the effect of the response groups on survival and the effect of the derived score at day 43 on survival. Kaplan–Meier curves were produced to visualize these effects. All analyses were performed using R version 3.4.1 statistical software (R Foundation for Statistical Computing, Vienna, Austria). As no definitive statistical comparison was performed we did not correct for multiple testing.

Results

Baseline characteristics

The baseline characteristics of the 51 patients included in the analysis are listed in Table 2 & Appendix Table A.1. No *EGFR* or *ALK* (ALK receptor tyrosine kinase) mutations were present within the cohort. The median follow-up was 538 days with a minimum follow-up of 341 days.

Response evaluation

Based on the changes in size of target lesions determined by CT scan, we were able to identify four distinct response groups (see methods for definitions): PsPr (n = 6), 2) OR (n = 10), 3) SD (n = 10) and 4) RPD (n = 25). Five of the 25 patients who were classified as RPD died after one CT scan but before the second scan due to lung cancer; however, they were still regarded as RPD (Figure 1 & Appendix Figure A.2). We did not find any patient with a hyperprogressive pattern.

Laboratory data

Relative changes to baseline in the laboratory values showed some correlation with response, as shown in Figure 2, & Appendix Figures A.3 & A.4. The CRP level changes over time correlated partially with radiological response to therapy. A case of upper respiratory infection led to a pronounced increase in the CRP level in one patient with PsPr, leading to a confounding increase in the median for the whole PsPr group. The LDH levels developed differently in the four response groups identified, with a continuously decreasing level for the OR group and a continuous increase for the RPD group. The SD and PsPr groups showed only minor changes in LDH over time. The NLR showed a continuous decrease for the PsPr and OR groups, as well as a steady increase for the RPD group, without any significant changes for the SD group (Figure 2).

Predictive score performance

After only four cycles and one CT scan the differences in laboratory values were small and mostly overlapping. Thus, in order to produce a more robust and earlier signal, we combined the values in a weighted score (Table 1, Figure 3 & Appendix Figure A.5). The ROC analyses led to a score combining baseline values for LDH, CRP, and the NLR, relative changes in LDH, CRP, and the NLR, and the change in tumor size on the first CT scan after commencing nivolumab treatment after four cycles of nivolumab at day 43 (Figure 4). See the score development section for details of score computation.

The score provided a sensitivity of 92.3% to detect those patients who actually benefitted from nivolumab after only four cycles (specificity 0.50, AUC = 0.80) (Figure 4). Prediction of benefit versus no benefit at day 43 led to a hazard ratio (HR) for OS of 0.38 (95% CI: 0.18-0.82; p =0.01) (Figure 5B).

The summarized scores within a group were compared with the RPD group as reference (Figure 3). All patients that derived benefit (SD, PsPr and OR groups) exhibited higher scores than those patients classified with RPD. The score significantly reflects the differences between the RPD and OR groups on days 43 and 99 ($p = 3.4 \times 10^{-5}$ and p = 0.00022, respectively). The median scores of the RPD and PsPr groups differed significantly on day 43, with a median score of 3.5 for the RPD group and 5.5 for the PsPr group (p = 0.0151). On day 99, the median score decreased to 3 for the RPD group and increased to 6 for the PsPr group (p = 0.020).

As the box plots shown in Figure 3 could only reflect a group of patients showing the same response, we also evaluated the score on a single-patient basis employing a cutoff of 5.5 (Appendix Figure A.5). With a cutoff of

Table 2. Baseline characteristics of the 51 patients included in the study analysis.							
Parameter	Objective response	Pseudoprogression	Real progressive disease	Stable disease	p-value [†]		
Patients, n	10	6	25	10			
Age, years					0.57		
$\text{Mean} \pm \text{SD}$	64 ± 11	68 ± 12	67 ± 11	70 ± 5.2			
Median (range)	61 (46–80)	72 (52–78)	66 (35–84)	68 (63–77)			
Age class, n (%)					0.51		
0–50 years	1 (10.0)	0 (0.0)	2 (8.0)	0 (0.0)			
51–60 years	3 (30.0)	2 (33.3)	3 (12.0)	0 (0.0)			
61–70 years	3 (30.0)	1 (16.7)	11 (44.0)	6 (60.0)			
\geq 70 years	3 (30.0)	3 (50.0)	9 (36.0)	4 (40.0)			
Sex, n (%)					0.03		
Female	8 (80.0)	3 (50.0)	7 (28.0)	3 (30.0)			
Male	2 (20.0)	3 (50.0)	18 (72.0)	7 (70.0)			
Histology, n (%)					0.20		
ADC	7 (70.0)	3 (50.0)	10 (40.0)	4 (40.0)			
LCNEC	0 (0.0)	1 (16.7)	0 (0.0)	0 (0.0)			
SCC	3 (30.0)	2 (33.3)	15 (60.0)	6 (60.0)			
Smoking status, n (%)					0.65		
Never	1 (10.0)	0 (0.0)	0 (0.0)	1 (10.0)			
Previous	6 (60.0)	4 (66.7)	17 (68.0)	5 (50.0)			
Current	3 (30.0)	2 (33.3)	8 (32.0)	4 (40.0)			
Stage, <i>n</i> (%)							
IIIB	0 (0.0)	0 (0.0)	1 (4.0)	0 (0.0)			
IV	10 (100.0)	6 (100.0)	24 (96.0)	10 (100.0)			
PD-L1 TPS, n (%)					0.48		
<1%	6 (60.0)	1 (16.7)	7 (28.0)	5 (50.0)			
≥1%	1 (10.0)	0 (0.0)	1 (4.0)	0 (0.0)			
>50%	1 (10.0)	0 (0.0)	2 (8.0)	1 (10.0)			
Missing	2	5	15	5			
PD-L1 IC, n (%)					0.03		
<1%	2 (20.0)	0 (0.0)	3 (12.0)	5 (50.0)			
≥1%	6 (60.0)	1 (16.7)	7 (28.0)	1 (10.0)			
>50%	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)			
Missing	2	5	15	4			
KRAS, n (%)					0.33		
Mutation exon 2	0 (0.0)	1 (50.0)	4 (57.1)	1 (50.0)			
Wild-type	4 (100.0)	1 (50.0)	3 (42.9)	1 (50.0)			
Missing	6	4	18	8			
EGFR, n (%)					1.00		
Wild-type	7 (100.0)	4 (100.0)	9 (100.0)	2 (100.0)			
Mutated	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)			
Missing	3	2	16	8			
Line, <i>n</i> (%)					0.79		
Second	6 (60.0)	3 (50.0)	14 (56.0)	6 (60.0)			

[†]Clinical parameters at baseline were tested for significant differences (p-value) between each response group. Fisher's exact test was used for all variables except 'age', which was tested by means of one-way ANOVA, and 'therapy duration', which contains censored cases and was tested using a likelihood ratio test between the Cox model 'therapy duration' and a null model.

[‡]Therapy duration refers to the date of 1 August 2017.

ADC: Adenocarcinoma; IC: Immune cells as percentage of area of immune infiltrate positive for PD-L1 in immunohistochemistry; LCNEC: Large cell neuroendocrine carcinoma; NA: Not available; SCC: Squamous cell carcinoma; SE: Standard error; TPS: Tumor proportion score.

Table 2. Baseline characteristics of the 51 patients included in the study analysis (cont.).								
Parameter	Objective response	Pseudoprogression	Real progressive disease	Stable disease	p-value [†]			
Third	3 (30.0)	2 (33.3)	8 (32.0)	2 (20.0)				
Fourth	0 (0.0)	1 (16.7)	3 (12.0)	2 (20.0)				
Fifth	1 (10.0)	0 (0.0)	0 (0.0)	0 (0.0)				
Side, <i>n</i> (%)					0.64			
Left	2 (22.2)	1 (16.7)	10 (41.7)	3 (30.0)				
Right	7 (77.8)	5 (83.3)	14 (58.3)	7 (70.0)				
Missing	1	0	1	0				
Therapy duration, weeks \ddagger					<0.001			
Median (95% CI)	72 (63.6–NA)	33.4 (24–NA)	16 (14–20.3)	36.6 (21.9–NA)				
Restricted mean \pm SE	75 ± 7.4	41.1 ± 10.5	17.7 ± 1.7	43.5 ± 7.6				

[†]Clinical parameters at baseline were tested for significant differences (p-value) between each response group. Fisher's exact test was used for all variables except 'age', which was tested by means of one-way ANOVA, and 'therapy duration', which contains censored cases and was tested using a likelihood ratio test between the Cox model 'therapy duration' and a null model.

[‡]Therapy duration refers to the date of 1 August 2017.

ADC: Adenocarcinoma; IC: Immune cells as percentage of area of immune infiltrate positive for PD-L1 in immunohistochemistry; LCNEC: Large cell neuroendocrine carcinoma; NA: Not available; SCC: Squamous cell carcinoma; SE: Standard error; TPS: Tumor proportion score.

5.5, nivolumab was predicted to be beneficial after four cycles (day 43) in all patients who, after eight cycles, were radiologically identified as OR (ten out of ten), five out of six patients with PsPr and eight out of ten patients with SD. Conversely, only four out of 25 patients in the RPD group were erroneously categorized as positive (Appendix Figure A.5).

Outcomes

The Kaplan–Meier survival analysis revealed a statistically significant difference in survival probability between the RPD and OR groups (HR: 0.13, 95% CI: 0.03–0.56; p < 0.01) (Figure 5A & Appendix Figure A.6). Half of the patients in the RPD group died less than 26 weeks after initiation of nivolumab treatment. At the time of last survival analysis (23 April 2019) only one patient in the OR group had died 1 year after treatment initiation with nivolumab. Nine patients (90%) were alive for more than 1 year compared to 70% (SD group), 66.7% (PsPr group) and 24% (RPD group) (Figure 5A).

A numerical (χ^2 = 3.63, p = 0.057) difference between squamous cell lung cancer and nonsquamous NSCLC was detected, with the latter tending to show more responses and less progressive disease (Appendix Tables A.2 & A.3, & Appendix Figure A.7).

None of the clinical parameters analyzed (e.g., fever, cough, shortness of breath, pain, general condition or change in weight over time; Appendix Figure A.3) showed any correlation with the different response groups.

Six out of 68 patients (8.8%) developed immune-related adverse events (AEs) grade 2 or higher that led to discontinuation of nivolumab and the initiation of prednisolone at 0.5–1.0 mg/kg. Four patients developed pneumonitis (7.8%), two patients developed colitis (3.9%) and one patient developed nephritis (2%). One of these patients developed pneumonitis and colitis simultaneously. All immune-related AEs could be reversed by steroids and none was fatal [20]. Additionally, five cases of grade 1 immune-related AEs were recorded that led to no change in therapy and required no systemic steroids: one case of arthralgia and myalgia (2%) and four mild cases of skin toxicity grade 1 (7.8%) that could be treated with topical steroids alone. Median time to onset of immune-related AEs was 21 weeks (3–88 weeks). Events leading to discontinuation of nivolumab occurred at a median of 56 weeks (6–88 weeks).

The AEs occurred independently of the efficacy of nivolumab and did not help predicting efficacy. Only three of any grade immune-related AEs occurred before week 6 (day 43) in two patients who had a best response of SD and one of RPD. All other immune-related AEs grade 2 or higher occurred at week 15 or later.

Radiological parameters failed to predict the efficacy of nivolumab apart from changes in size. Neither blurred appearance, central necrosis nor any other radiological finding correlated with efficacy. In particular, we could not find different radiological patterns to visually distinguish atypical responses from RPD. Relevant radiographs are presented in Appendix Figures A.8–A.14. In addition, baseline characteristics of the response groups did not show any significant differences (Appendix Table A.1).



Figure 2. Relative changes over time. Relative changes of **(A–C)** biomarkers and **(B)** size of target lesions in each response group over time. Baseline (day 1) = 0%. Lines connect median values of the box plots. G/L: Granulocyte (neutrophil)/lymphocyte.

Discussion

In a retrospective analysis, we were able to create a weighted score based on common laboratory parameters (LDH, CRP and the NLR) that allowed us to distinguish individual patients deriving benefit from nivolumab therapy from those who do not with a sensitivity of 92.3%. Prediction of benefit after only four cycles of therapy with nivolumab led to a HR of 0.31 (p < 0.01).

Currently, discussed predictive tests such as tumor mutational burden lead to a HR for OS of 0.64 (95% CI: 0.44–0.93) [21] with the disadvantage that with a higher threshold the HR can be improved by enrichment but simultaneously more patients who would have benefitted are excluded from CIT. As no reliable predictive marker is currently available most patients with NSCLC will be treated with CIT eventually, necessitating means of monitoring efficacy concurrent with therapy.

The standard procedure to monitor efficacy of oncological treatment is radiological evaluation by repeated CT scans. However, in CIT, the phenomenon of atypical and delayed responses can lead to equivocal radiological results. Accordingly, attempts to deal with atypical responses have led to modifications of classical RECIST version 1.1 (e.g., iRECIST) with an increase in response rates of only 1–2% [18,19]. Nevertheless, to draw definite conclusions according to RECIST as well as iRECIST requires a second CT scan, delaying the decision for an additional



Figure 3. Predicted scores. Predicted scores on **(A)** day 43 and **(B)** day 99 shown separately for each response group. The p-values reflect the comparisons against the real progressive disease group using the Mann–Whitney *U*-test (Holm adjusted).

6–8 weeks. Our score, however, was able to predict benefit after only one CT scan, leading to an earlier decision to proceed with CIT or withdraw therapy. It seems important to detect patients who derive no benefit from CIT and who should receive alternative treatment as early as possible. Sorting patients with our score after 6 weeks of treatment and with only one CT scan showed a survival of only 24% of patients after 1 year when predicted as not benefitting compared with 73% of patients after 1 year who were predicted to benefit.

To the best of our knowledge, this is the first published research addressing the value of relative changes of multiple laboratory values over time in immuno-oncology. The laboratory parameters we selected had previously been described as prognostic or predictive markers in cancer therapy. A NLR greater than 2, CRP greater than 0.5 mg/dl, as well as high LDH are considered to be linked to an unfavorable prognosis [20,22–26].

future science group fsg



Figure 4. Receiver operating characteristic curve analyses. Solid gray lines show the performance in each repeat; the solid black line is the consensus ROC curve across all ten repeats. The Youden index of the consensus ROC curve is marked with a dot. The AUC values of the consensus ROC curve as well as sensitivity and specificity at the Youden index are indicated in the lower right corner of the plots. G/L: Granulocyte (neutrophil)/lymphocyte; ROC: Receiver operating characteristic.



Figure 5. Survival analysis. (A) Survival analysis in different radiological response groups according to RECIST version 1.1 using Kaplan–Meier curves for each response group. **(B)** Survival analysis in different response groups predicted at day 43 by the proposed score. Those patients benefitting from nivolumab were amalgamated as 'no RPD' (i.e., OR, PsPr and SD) versus those patients predicted not to benefit (i.e., RPD). The two Kaplan–Meier curves represent samples that the predictor classified 'no RPD'. The tables below the plots show the patients at risk.

OR: Objective response; PsPr: Pseudoregression; RPD: Real progressive disease; SD: Stable disease.

Recently, Bagley *et al.* [27] found the NLR at baseline to be a marker predictive of response to nivolumab using a cut-off value of 5. The results of our patient cohort did not support a NLR cutoff of 5 as we found five ORs each in patients with a NLR below 5 as well as above 5. In the RPD group, we found 12 patients to have a baseline NLR above 5 and 13 patients to have a baseline NLR below 5. However, we could show that relative changes of NLR were of some value as the index further increased in patients with progressive disease and normalized in all other response groups. The difference might be explained because Bagley *et al.* [27] did not use standardized RECIST evaluation, whereas all our patients were evaluated according to RECIST version 1.1 and iRECIST [19].

Furthermore, we were able to show that PsPr is a rather common feature with an incidence of 12%, which is higher than in most other case series of patients treated with nivolumab [28,29]. The frequency of PsPr is probably



Figure 5. Survival analysis (cont.). (A) Survival analysis in different radiological response groups according to RECIST version 1.1 using Kaplan–Meier curves for each response group. **(B)** Survival analysis in different response groups predicted at day 43 by the proposed score. Those patients benefitting from nivolumab were amalgamated as 'no RPD' (i.e., OR, PsPr and SD) versus those patients predicted not to benefit (i.e., RPD). The two Kaplan–Meier curves represent samples that the predictor classified 'RPD' and samples that the predictor classified 'no RPD'. The tables below the plots show the patients at risk.

OR: Objective response; PsPr: Pseudoregression; RPD: Real progressive disease; SD: Stable disease.

dependent on the interval chosen for CT controls as we were also able to demonstrate early PsPr that was only visible on chest radiographs obtained at the second to fourth cycle of nivolumab therapy, but that had already turned into a partial remission or SD by the time of the first CT control (Appendix Figures A.8–A.14).

Given the small numbers (n = 6 in the PsPr group), the difference in our score between the PsPr and RPD groups on day 43 did not reach statistical significance (p = 0.051), but seems to be particularly noteworthy. Most clinicians would continue CIT with no hesitation in patients with SD without any sign of tumor shrinkage. However, a score to guide the decision in patients with radiological signs of tumor progression would be extremely helpful. Only a single patient classified as PsPr after two CT scans showed a score below 5.5 at the first CT control. Thus, despite a gain in tumor size in the corresponding CT scan, the score indicated a benefit from nivolumab in five of the six patients who went on to develop tumor shrinkage, whereas those patients who would eventually develop progressive disease showed a score below 5.5 in 75% of cases (Appendix Figure A.5).

Clinical parameters such as weight loss, AEs or radiological signs were not found to be related to response. Some studies have already attempted to detect PsPr using positron emission tomography-CT or magnetic resonance imaging. The results indicate that these techniques have not yet been able to conclusively distinguish between PsPr, RPD and tumor necrosis, especially after only a few cycles of CIT [30–32]. The fact that even experienced radiologists are not able to find valid radiological signs to reveal PsPr supports the importance of our data.

Some investigators have reported that immune-related AEs could be used as a monitoring tool in patients treated with immune checkpoint inhibitors, for example, Haratani *et al.* [33]. However, the publication by Haratani *et al.* [34] lacks one essential piece of information: the duration of treatment. A recent publication by de Marinis *et al.* [34] concluded that immune-related AEs were not correlated with response but with treatment exposure, in other words, those patients who derive benefit from therapy will be on CIT for longer and thus develop immune-related AEs more frequently. In the 2017 European Society for Medical Oncology guidelines for immune-related AEs, the median time to onset of immune-related AEs in relation to different organs is given as: skin toxicity 19.4 weeks, gastrointestinal toxicity 26.3 weeks, endocrine 28.6 weeks and nephritis even up to 50.9 weeks [35]. Thus, one would obviously miss most of these immune-related AEs when patients are not deemed responsive and taken off immune checkpoint inhibitors after only four cycles (6 weeks). We could not find any correlation of clinical benefit with immune-related AEs in our cohort as most patients did not develop immune-related AEs before week 15, which is in our opinion too late to decide on the benefit of CIT treatment. Only three patients developed immune-related AEs during the first 6 weeks of treatment, of which one was found to have RPD and two were found to have SD. None of the patients with PsPr or OR developed immune-related AEs early on.

We did not treat any patient with a targetable driver mutation in our cohort as, at least for patients with an *EGFR* mutation, checkpoint inhibitors seem to be less effective. As nivolumab was approved in the European Union for squamous cell lung cancer in October 2015 and for nonsquamous NSCLC in April 2016, our cohort included a larger number of squamous cell lung cancer patients than would be expected, thus leaving only a few patients with nonsquamous NSCLC and even fewer patients with a detectable *KRAS* mutation [36,37]. Of these two showed RPD and one SD with real progression after 8 months of nivolumab therapy; thus, the previously observed good performance of *KRAS*-mutated patients treated with immune checkpoint inhibitors could not be reproduced in our study. Since this was a retrospective study, in many cases there was not enough tissue left to perform further pathological diagnostics, such as determining the tumor mutational burden.

As PD-L1 testing could only be provided for a fraction of patients we were not able to draw any statistical conclusions regarding the correlation of our score with PD-L1 immunohistochemistry. However, we found at least one patient without any expression of PD-L1 and an immediate, still ongoing remission. Other response patterns (e.g., hyperprogression) were not seen in our study [38].

The most obvious limitation of our trial is the small number of patients and the fact that we performed this analysis in a single center. As a next step, it would be very interesting to validate our score with the large datasets aggregated in Phase III clinical trials of CIT. However, as this was a single-center trial the study was performed in a rather homogeneous cohort with no patient lost to follow-up. We also acknowledge that the only meaningful response assessment in a single-arm retrospective study is response; however, we were interested to find that real survival is linked to our proposed scoring tool.

Conclusion

We developed a weighted score combining three common laboratory parameters (LDH, CRP and the NLR) and the first CT scan under therapy that could help to predict patients who will benefit from CIT after only 6 weeks of treatment. Larger datasets are needed to validate our findings.

Supplementary data & appendix

To view the supplementary data that accompany this paper please visit the journal website at: www.futurescience.com/doi/suppl/10.2217/imt-2019-0039

Acknowledgments

We thank the patients and their families, as well as the staff of the LKI Immenhausen (Immenhausen, Germany) who cared for the patients, for making this trial possible. Medical writing assistance was provided by Ray Loughlin (Academic Editor, Sheffield, UK).

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

The authors state that they have obtained appropriate institutional review board approval or have followed the principles outlined in the Declaration of Helsinki for all human or animal experimental investigations. In addition, for investigations involving human subjects, informed consent has been obtained from the participants involved.

Author contributions

A Schiwitza, S Andreas and A Rittmeyer focused on the conception and design. All authors were dedicated to the collection and assembly of data, data analysis and interpretation, manuscript writing, final approval of manuscript. All authors are accountable for all aspects of the work.

Open access

CC-BY-NC-ND: this work is licensed under the Attribution-NonCommercial-NoDerivatives 4.0 Unported License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/4.0/

Summary points

- Atypical and delayed responses are a common feature in patients treated with checkpoint inhibitors and predictive tests are not very reliable.
- Combining a few routinely collected laboratory values (LDH, CRP and the NLR) and a first CT scan after only 6 weeks of therapy is able to predict long-term benefit from therapy with checkpoint inhibitors.
- Integrating these laboratory parameters in a weighted score has superior predictive power compared with single parameters alone.
- Using this approach enables patients with pseudoprogression to be reliably distinguished from patients with real progressive disease.
- This study provides for the first time an attempt to monitor efficacy of checkpoint inhibitors concurrent with treatment and to detect patients who benefit during ongoing CIT.

References

- 1. Ferlay J, Soerjomataram I, Dikshit R *et al.* Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer.* 136(5), E359–E386 (2015).
- 2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. CA Cancer J. Clin. 66(1), 7-30 (2016).
- 3. Rittmeyer A, Barlesi F, Waterkamp D *et al.* Atezolizumab versus docetaxel in patients with previously treated non-small-cell lung cancer (OAK): a Phase III, open-label, multicentre randomised controlled trial. *Lancet* 389(10066), 255–265 (2017).
- Brahmer J, Reckamp KL, Baas P et al. Nivolumab versus docetaxel in advanced squamous-cell non-small-cell lung cancer. N. Engl. J. Med. 373(2), 123–135 (2015).
- Borghaei H, Paz-Ares L, Horn L et al. Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. N. Engl. J. Med. 373(17), 1627–1639 (2015).
- Herbst RS, Baas P, Kim D-W *et al.* Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. *Lancet* 387(100027), 1540–1550 (2016).
- Reck M, Rodríguez-Abreu D, Robinson AG *et al.* Pembrolizumab versus chemotherapy for PD-L1-positive non-small-cell lung cancer. N. Engl. J. Med. 375(19), 1823–1833 (2016).
- Gettinger SN, Horn L, Gandhi L *et al.* Overall survival and long-term safety of nivolumab (antiprogrammed death 1 antibody, BMS-936558, ONO-4538) in patients with previously treated advanced non-small-cell lung cancer. *J. Clin. Oncol.* 33(18), 2004–2012 (2015).
- 9. Brahmer JR, Tykodi SS, Chow LQM *et al.* Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. *N. Engl. J. Med.* 366(26), 2455–2465 (2012).
- Topalian SL, Hodi FS, Brahmer JR *et al.* Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N. Engl. J. Med.* 366(26), 2443–2454 (2012).

- 11. Sorensen SF, Zhou W, Dolled-Filhart M et al. PD-L1 expression and survival among patients with advanced non-small cell lung cancer treated with chemotherapy. *Transl. Oncol.* 9(1), 64–69 (2016).
- 12. Tanaka Y, Morita CT, Okamura H. Anti-PD-1 and anti-PD-L1 mAbs. In: Immunotherapy of Cancer: An Innovative Treatment Comes of Age. Yamaguchi Y (Ed.). Springer, Tokyo, Japan, 383–294 (2016).
- Sundar R, Cho B-C, Brahmer JR, Soo RA. Nivolumab in NSCLC: latest evidence and clinical potential. *Ther. Adv. Med. Oncol.* 7(2), 85–96 (2015).
- Rizvi NA, Mazières J, Planchard D *et al.* Activity and safety of nivolumab, an anti-PD-1 immune checkpoint inhibitor, for patients with advanced, refractory squamous non-small-cell lung cancer (CheckMate 063): a Phase II, single-arm trial. *Lancet Oncol.* 16(3), 257–265 (2015).
- 15. Chiou VL, Burotto M. Pseudoprogression and immune-related response in solid tumors. J. Clin. Oncol. 33(31), 3541–3543 (2015).
- Kerr KM, Tsao M-S, Nicholson AG et al. Programmed death-ligand 1 immunohistochemistry in lung cancer: in what state is this art? J. Thorac. Oncol. 10(7), 985–989 (2015).
- 17. Kolla BC, Patel MR. Recurrent pleural effusions and cardiac tamponade as possible manifestations of pseudoprogression associated with nivolumab therapy a report of two cases. *J. Immunother. Cancer* 4, 80 (2016).
- 18. Hodi FS, Ballinger M, Lyons B et al. Immune-modified response evaluation criteria in solid tumors (imRECIST): refining guidelines to assess the clinical benefit of cancer immunotherapy. J. Clin. Oncol. 36(9), 850–858 (2018).
- Seymour L, Bogaerts J, Perrone A et al. iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. Lancet Oncol. 18(3), e143–e152 (2017).
- Abdel-Wahab N, Shah M, Suarez-Almazor ME. Adverse events associated with immune checkpoint blockade in patients with cancer: a systematic review of case reports. *PLoS ONE* 11(7), e0160221 (2016).
- Gandara DR, Kowanetz M, Mok TSK *et al.* Blood-based biomarkers for cancer immunotherapy: tumor mutational burden in blood (bTMB) is associated with improved atezolizumab (atezo) efficacy in 2L+ NSCLC (POPLAR and OAK). *Nat Med.* 24(9), 1441–1448 (2018). doi: 10.1038/s41591-018-0134-3
- 22. Yamaguchi Y (Ed.). Kawakami personalized cancer immunotherapy (Chapter 24). In: *Immunotherapy of Cancer: An Innovative Treatment Comes of Age.* Tokyo, Japan, Springer (2016).
- 23. Lee S, Choe J-W, Kim H-K, Sung J. High-sensitivity C-reactive protein and cancer. J. Epidemiol. 21(3), 161-168 (2011).
- 24. Aref H, Refaat S. CRP evaluation in non-small cell lung cancer. Egypt J. Chest Dis. Tuberc. 63(3), 717-722 (2014).
- 25. Yu S-L, Xu L-T, Qi Q et al. Serum lactate dehydrogenase predicts prognosis and correlates with systemic inflammatory response in patients with advanced pancreatic cancer after gemcitabine-based chemotherapy. Sci. Rep. 7, 45194 (2017).
- Gopalakrishnan V, Spencer CN, Nezi L et al. Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. Science 359(6371), 97–103 (2018).
- 27. Bagley SJ, Kothari S, Aggarwal C *et al.* Pretreatment neutrophil-to-lymphocyte ratio as a marker of outcomes in nivolumab-treated patients with advanced non-small-cell lung cancer. *Lung Cancer* 106, 1–7 (2017).
- Cancerworld. The role of immunotherapy in treating solid cancers. European School of Oncology (2017). http://cancerworld.net/e-grandround/the-role-of-immunotherapy-in-treating-solid-cancers
- 29. Deel A. Nivolumab in metastatic non-small cell lung cancer. J. Adv. Pract. Oncol. 7(2), 220-225 (2016).
- Nasseri M, Gahramanov S, Netto JP *et al.* Evaluation of pseudoprogression in patients with glioblastoma multiforme using dynamic magnetic resonance imaging with ferumoxytol calls RANO criteria into question. *Neuro Oncol.* 16(8), 1146–1154 (2014).
- 31. Parvez K, Parvez A, Zadeh G. The diagnosis and treatment of pseudoprogression, radiation necrosis and brain tumor recurrence. *Int. J. Mol. Sci.* 15(7), 11832–11846 (2014).
- 32. Aguiar PN Jr, Santoro IL, Tadokoro H *et al.* A pooled analysis of nivolumab for the treatment of advanced non-small-cell lung cancer and the role of PD-L1 as a predictive biomarker. *Immunotherapy* 8(9), 1011–1019 (2016).
- Haratani K, Hayashi H, Chiba Y et al. Association of immune-related adverse events with nivolumab efficacy in non-small-cell lung cancer. JAMA Oncol. 4(3), 374–378 (2018).
- 34. de Marinis F, Barlesi F, Rittmeyer A *et al.* Survival and safety of atezolizumab by best overall response (BOR) in the Phase III NSCLC OAK study. *Ann. Oncol.* 28(Suppl 5), v460–v496 (2017).
- 35. Haanen JBAG, Carbonnel F, Robert C *et al.* Management of toxicities from immunotherapy: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* 28(Suppl 4), iv119–iv142 (2017).
- Kempf E, Rousseau B, Besse B, Paz-Ares L. KRAS oncogene in lung cancer: focus on molecularly driven clinical trials. *Eur. Respir.* Rev. 25(139), 71–76 (2016).
- Rangachari D, VanderLaan PA, Shea M *et al.* Correlation between classic driver oncogene mutations in *EGFR*, *ALK*, or *ROS1* and 22C3-PD-L1 ≥50% expression in lung adenocarcinoma. *J. Thorac. Oncol.* 12(5), 878–883 (2017).
- Champiat S, Dercle L, Ammari S et al. Hyperprogressive disease is a new pattern of progression in cancer patients treated by anti-PD-1/PD-L1. Clin. Cancer Res. 23(8), 1920–1928 (2017).