Comparison between 18F-FDG-PET- and CT-based criteria in non-small cell lung cancer (NSCLC) patients treated with Nivolumab

Comparison Between ^{18}F-FDG PET–Based and CT-Based Criteria in Non–Small Cell Lung Cancer Patients Treated with Nivolumab

Giovanni Rossi[1,2,] Matteo Bauckneht[3], Carlo Genova[1], Erika Rijavec[4], Federica Biello[1], Simone Mennella[5], Maria Giovanna Dal Bello[1], Giuseppe Cittadini[5], Paolo Bruzzi[6], Roberta Piva[3], Valentina Ceriani[3], Gianmario Sambuceti[3,7], Egesta Lopci[8], Silvia Morbelli[3,7], and Francesco Grossi[4]

[1]Lung Cancer Unit, IRCCS Policlinico San Martino, Genoa, Italy;

[2]Department of Clinical, Surgical, and Experimental Sciences, Division of Experimental Pathology and Oncology, University of Sassari, Sassari, Italy;

[3]Nuclear Medicine Unit, IRCCS Policlinico San Martino, Genoa, Italy;

[4]Medical Oncology Unit, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy;

[5]Radiology Unit, IRCCS Policlinico San Martino, Genoa, Italy;

[6]Epidemiology Unit, IRCCS Policlinico San Martino, Genoa, Italy;

[7]Department of Health Sciences (DISSAL), University of Genoa, Genoa, Italy;

[8]Nuclear Medicine Unit, Humanitas Clinical and Research Center–IRCCS, Rozzano, Italy

For correspondence contact: Silvia Morbelli, Nuclear Medicine Unit, Department of Health Sciences (DISSAL), University of Genoa, IRCCS Policlinico San Martino, Largo R. Benzi 10, Genoa, 16132, Italy. E-mail: silvia.morbelli@unige.it

Abstract

Because of the peculiar mechanism of action of immune checkpoint inhibitors (ICIs), evaluation of the radiologic response to them in solid tumors presents many challenges. We aimed to compare evaluation of the first response to nivolumab by means of CT-based criteria with respect to ^{18}F-FDG PET response criteria in non–small cell lung cancer (NSCLC) patients.

Methods:

Seventy-two patients with advanced NSCLC were recruited in a single-institution ancillary trial within the expanded-access program (NCT02475382) for nivolumab. Patients underwent CT and ^{18}F-FDG PET at baseline and after 4 cycles (the first evaluation). In cases of progressive disease, an additional evaluation was performed after 2 further cycles to confirm progression. We evaluated the treatment response on CT using RECIST 1.1 and the immune-related response criteria (irRC) and on ^{18}F-FDG PET using PERCIST and immunotherapy-modified PERCIST. The concordance between CT- and PET-based criteria and the capability of each method to predict overall survival were evaluated.

Results:

Forty-eight of 72 patients were evaluable for a first response assessment with both PET- and CT-based criteria. We observed low concordance between CT- and PET-based criteria (κ-value of 0.346 and 0.355 between PERCIST and imPERCIST and RECIST, respectively. κ-value of 0.128 and 0.198 between PERCIST and imPERCIST and irRC, respectively). Regarding overall survival, irRC could more reliably distinguish responders from nonresponders. However, thanks to the prognostic value of partial metabolic response assessed by both PERCIST and immunotherapy-modified PERCIST, PET-based response maintained prognostic significance in patients classified as having progressive disease on the basis of irRC.

Conclusion:

Even though the present study did not support the routine use of ^{18}F-FDG PET in the general population of NSCLC patients treated with ICIs, the findings suggest that metabolic response assessment has added prognostic value, potentially improving therapeutic decision making.

Key Words: NSCLC; checkpoint inhibitors; PET; CT

Response to therapy in solid tumors is conventionally monitored by morphologic imaging. Traditionally, tumor shrinkage describes treatment success, and if not seen, patients are assumed to be nonresponders. The 2 most widely used systems for the classification of tumor shrinkage, RECIST and the system proposed by the World Health Organization, were developed to standardize response evaluation in phase II clinical trials (**1**). However, although patients who respond to treatment are known to have a better prognosis, the validity of an objective response to chemotherapy as a surrogate endpoint of survival is more controversial, especially in some clinical settings (**2**).

In recent years, the problem became even more pronounced with the introduction of targeted anticancer therapies and for all regimens with immune checkpoints inhibitors (ICIs) (**2**,**3**). Indeed, the anticancer immune reaction activated by ICIs may initially increase the total tumor volume because of inflammatory cell infiltrates that mimic cancer progression (**4**). Therefore, atypical response patterns termed pseudoprogression might be observed in patients who receive ICIs. These patients may initially meet the conventional response criteria for progressive disease (PD) but later show a reduction in tumor burden (**5**,**6**); hence, discontinuation of treatment on the basis of disease progression as defined according to RECIST might be premature. Therefore, clinical trials for ICIs often allow treatment beyond RECIST-defined progression (**1**). On the other hand, RECIST has been adapted to overcome this limitation by the creation of immune-related response criteria (irRC) (**7**) and, more recently, immune-RECIST (**8**).

In this scenario, although the morphology-based criteria are being further validated, less evidence is available on the added value of ^{18}F-FDG PET in patients treated with ICIs (**9**). Despite the potential occurrence of inflammatory infiltration and related tumor changes that might also hamper the reliability of the PET signal, ^{18}F-FDG PET might be able to capture the response to ICIs in specific subgroups of patients. However, the exact positioning of ^{18}F-FDG PET in the flowchart of patients treated with ICIs and its cost-effectiveness with respect to conventional morphologic assessment still need to be defined. Moreover, PERCIST was introduced in 2009 as a guideline for the structured ^{18}F-FDG PET assessment of response to oncologic therapy, but the frequency and impact of pseudoprogression as seen by PERCIST in patients treated with ICIs is not well documented. Similarly, immunotherapy-modified PERCIST (imPERCIST) has been proposed but is not yet fully validated in patients with melanoma treated with ipilimumab, and even less evidence is available for the added value of PET-based response in non–small cell lung cancer (NSCLC) patients treated with ICIs (**10**–**13**). The aim of the present study was to assess, in patients with advanced NSCLC, the frequency and pattern of the ^{18}F-FDG PET–based first response to nivolumab, a fully human IgG4 program-death-1 antibody (**14**), and to compare them with both RECIST 1.1 and the CT-based irRC. To this aim, PET response was evaluated with both standard PERCIST and imPERCIST (**15**). As a secondary aim, the correlation between PET- or CT-based criteria and patients' overall survival (OS) was evaluated.

MATERIALS AND METHODS

Patients and Study Design

Seventy-four patients with advanced pretreated NSCLC were enrolled in a translational research trial at the Lung Cancer Unit of the IRCCS Policlinico San Martino. The trial was an ancillary single-institution study conducted within the expanded-access program for nivolumab (NCT02475382). Accordingly, the specific study design was approved by the research committee of Regione Liguria, and only patients enrolled at IRCCS Policlinico San Martino were included in the study. All enrolled patients gave written informed consent to participate in the study. Nivolumab was provided by Bristol-Meyers Squibb within the expanded-access program in NSCLC. The major inclusion criteria were an age of at least 18 y, histologically or cytologically confirmed NSCLC, a clinical stage of IIIb or IV (according to TNM, version 7.0), at least one previous line of therapy, at least one measurable lesion by RECIST 1.1, and previously treated or stable brain metastases from at least 2 wk before the treatment with nivolumab that did not need treatment with more than 10 mg/d of prednisone or the equivalent of another steroid. The exclusion criteria were a performance status of at least 3 on the Eastern Cooperative Oncology Group scale, meningeal carcinomatosis, active autoimmune disease or a syndrome requiring daily steroid treatment (except patients with diabetes mellitus type I and hypothyroidism requiring only hormone replacement), a previous line of therapy with ICIs, and the administration of a live attenuated vaccine within the 30 d before the first nivolumab administration.
CT and ^{18}F-FDG PET were performed as detailed previously (**16**) and in the "Image Acquisition Protocol" section below within 30 d before starting therapy with nivolumab, 3 mg/kg every 14 d. Imaging was repeated after 4 cycles (the first evaluation) and then every 4 cycles. Response to treatment was evaluated by CT using RECIST 1.1 (**3**) (hereafter referred to as RECIST) and irRC (**7**) and by ^{18}F-FDG PET using PERCIST (**15**). If patients experienced PD by RECIST, the protocol required that CT and ^{18}F-FDG PET be repeated after 2 additional cycles to confirm PD (irRC) and that patients be treated beyond progression in cases of clinical benefit. For patients with stable disease (SD) or a partial response (PR), CT and ^{18}F-FDG PET were repeated every 4 cycles (a schematic representation of the study design is shown in **Fig. 1**). Only patients who had at least one posttherapy evaluation with both CT and ^{18}F-FDG PET were included in the analyses; thus, early dropouts (i.e., due to early deaths) were not included in the analyses if at least one posttherapy evaluation (either CT or ^{18}F-FDG PET) was not available.

Image Acquisition Protocol

The CT parameters were as follows. For the arterial phase, slice thickness was 5 mm, pitch was 0.8, tube rotation speed was 0.5 s, voltage was 120 kV, reference 175 mA. A dose modulation system was applied to optimize total exposure according to the patient's body size; an additional set of 1-mm-thick slices was reconstructed to obtain high-resolution, multiplanar reformations. For the portal phase, slice thickness was 5 mm, pitch was 0.8, tube rotation speed was 0.5 s, voltage was 120 kV, reference 175 mA, with the same modulation system. Slices 2 mm thick at 1.5-mm intervals were reconstructed for multiplanar reformations. Iodinated contrast medium, with a concentration of 350 mg/mL, was injected

using a power injector at a flow rate of 3 mL/s and a dose of 80–130 mL, depending on body weight, followed by 40 mL of saline at the same flow rate. Standard 5-mm-thick images were used for rapid evaluation by the radiologist and for review by the referring physician, whereas thinner slices were used for multiplanar imaging of vessels and bone (ribs and spine) and for high-resolution scanning of lung and liver lesions.

^{18}F-FDG PET was performed according to international guidelines (**1**) using a 16-slice PET/CT hybrid system (Biograph 16; Siemens Medical Solutions). Briefly, patients fasted overnight before the intravenous administration of 300–400 MBq of ^{18}F-FDG, which was performed in a quiet room with the patient recumbent and still. Blood glucose was measured before tracer injection to confirm a level below 160 mg/dL. To minimize artifacts caused by activity in the urinary tract, patients were asked to drink 500 mL of water 1 h before image acquisition and to empty the bladder just before the acquisition began. Imaging started 60 ± 15 min after intravenous tracer administration. The technical parameters of the 16-detector helical CT scanner included a gantry rotation speed of 0.5 s and a table speed of 24 mm per gantry rotation. The PET component of the combined imaging system had an axial view of 16.2 cm per bed position, with an interslice space of 3.75 mm. The transaxial field of view and pixel size of the reconstructed PET images were 58.5 cm and 4.57 mm, respectively, with a matrix size of 128 × 128. Unenhanced low-dose CT was performed at 140 kV and 40 mA for attenuation correction of emission data and for anatomic localization of the PET dataset. The emission scan was performed in 3-dimensional mode shortly after the CT acquisition, with a 3-min acquisition per bed position. PET sinograms were reconstructed by iterative ordered-subset expectation maximization (3 iterations, 8 subsets). Scanning was performed starting from the orbital plane and continuing to the mid thigh, except for cases in which the clinical history demanded a whole-body, vertex-to-toes scan.

Definition of Group Response

CT findings were interpreted by physicians experienced in response evaluation with both RECIST and irRC, masked to the PET/CT results. Similarly, ^{18}F-FDG PET findings were interpreted according to standard PERCIST and imPERCIST by 2 nuclear medicine physicians experienced in PERCIST-based response evaluation, masked to the CT results. The response criteria have been detailed elsewhere (**3**,**7**,**15**).

Statistical Analysis

Patients without at least one CT scan and one ^{18}F-FDG PET scan before the initiation of nivolumab therapy and during it were excluded. Therefore, the proportion of objective responses and the OS reported in this study cannot be compared with those of other prognostic or therapeutic studies, because early deaths and dropouts had not undergone a first evaluation by means of ^{18}F-FDG PET and were excluded. Accordingly, survival curves start at 2 mo after the first treatment administration. The concordance between CT-based criteria and PERCIST was investigated by computing a Cohen κ-coefficient. Unweighted κ-values are reported using the benchmarks of Landis and Koch (**17**,**18**). Concordance between the first evaluation by RECIST and by irRC versus PERCIST and imPERCIST was calculated. To learn the prognostic value of treatment response assessed by ^{18}F-FDG PET alone or in addition to CT imaging, univariate OS curves by PERCIST response were computed

according to Kaplan–Meier analysis and compared using the log-rank test. To evaluate the relative contribution of each of the 3 classification systems to prognosis, a multivariate Cox proportional-hazards model was fitted to the data, with OS as the dependent variable and RECIST, irRC, PERCIST, and imPERCIST response as covariates. The final model was derived by means of a stepwise backward procedure based on the likelihood ratio test. For exploratory purposes, the possibility of a synergy between the imaging and metabolic techniques in improving the prognostic ability of the model was also evaluated by including in the model the appropriate interaction terms and by evaluating the resulting modification of its likelihood. All tests were 2-sided. Analyses were conducted with SPSS (release 23; IBM).

RESULTS

Of 74 patients included in the expanded-access program, 48 underwent both CT and ^{18}F-FDG PET at baseline and during their treatment course and had lesions characterized by a metabolism suitable for evaluation by PERCIST. Therefore, only these 48 patients were included in the present analysis (**Fig. 2**). The main characteristics of the patients are summarized in **Table 1**. At the first CT evaluation, 27 of 48 (56%) patients were classified as PD, 17 (36%) as SD, and 4 (8%) as PR according to RECIST. Conversely, irRC classified 24 (50%) patients as PD, 21 (44%) as SD, and 3 (6%) as PR. Finally, with PERCIST, 1 patient (2%) was classified as having a complete metabolic response, 11 (22%) as having a partial metabolic response (PMR), 22 (46%) as having stable metabolic disease (SMD), and 14 (29%) as having progressive metabolic disease (PMD). Classification according to imPERCIST substantially overlapped that according to PERCIST, with only 2 patients considered PMD according to PERCIST and conversely classified as SMD according to imPERCIST (in the entire cohort, imPERCIST classified 1 patient [2%] as having a complete metabolic response, 11 [22%] as PMR, 24 [50%] as SMD, and 12 [25%] as PMD).

Overall, the agreement between the first response evaluated by RECIST and both PET-based criteria was moderate, as only 58% of the patients were similarly classified by the 2 evaluations (κ-value of 0.346 with respect to PERCIST and 0.355 with respect to imPERCIST, $P = 0.001$; **Tables 2tbl2** and **3**). Of the 27 patients classified as PD by RECIST, 11 and 12 were classified as SMD by PERCIST and imPERCIST, respectively, and 2 as PMR by both PET-based criteria; by contrast, of the 15 patients classified as PMD by PERCIST, 1 was classified as SMD and none as PR by RECIST. Notably, the concordance between irRC on one side and PERCIST or imPERCIST on the other was much weaker ($\kappa =$ 0.128 and 0.198, $P = 0.218$; **Tables 4** and **5**). As was predictable, the 2 patients classified as SMD rather than PMD according to imPERCIST were also classified as SD according to irRC. Two representative examples of the mismatch between CT-based and PET-based criteria, as well as between PERCIST and imPERCIST, are shown in **Figure 3**.

In consideration of the moderate to low overall concordance between the different methods, subsequent analyses were performed to learn which method could be more reliable in predicting the prognosis and if any improvement could be achieved by combining them. The mean OS within the whole population is represented in **Figure 4**. On a per-criterion analysis, OS appeared to be similarly assessed by CT-based and PERCIST-based criteria (**Figs. 3A–3C**). Patients classified as PD according to RECIST, irRC, PERCIST, and imPERCIST showed a uniformly poor prognosis, with a median OS of 8.9, 8.4, 9.3, and 9.9 mo, respectively. However, differences were noticed among patients classified as SD. Indeed, in RECIST SD and irRC SD patients, OS was similar to or even better than that in patients

achieving PR. By contrast, in SMD patients, OS closely resembled that of patients with PMD (regardless of which PET-based approach was used). Finally, longer OS was predicted by PERCIST and imPERCIST PMR with respect to the few patients classified as RECIST PR and irRC PR (**Figs. 5A–5C**).

To examine more thoroughly the potential complementary role of CT- and PET-based methods to assess response, a multivariable Cox model was fitted to the data, with OS as a dependent variable and RECIST, irRC, or PERCIST classes. Because of the small numbers involved, a binary classification of response was used, with patients being categorized as simply responders or nonresponders for each of the 3 criteria. In the framework of clinical trials for advanced NSCLC, the CT-based disease control rate at the first response (CR + PR + SD) has demonstrated a significant positive predictive value (**19**). In the era of biologic targeted therapies and ICIs, it has been further demonstrated that the disease-control-rate metric more closely mirrors treatment effect than does the traditional response rate. Accordingly, in this post hoc analysis we classified all patients in PR or SD according to CT criteria as responders. Given the peculiar behavior of response to ICIs in previous studies, not only PMR but SMD as assessed by PET has also been considered indicative of response (**12**,**20**). However, in the present study, the inspection of the Kaplan–Meier curves highlighted a lower prognostic value for SMD, whose curves largely overlapped PMD curves both for PERCIST and imPERCIST. Accordingly, to fully exploit the residual prognostic value of PET-based response in irRC patients, we decided to evaluate the residual prognostic role of PET in irRC response categories considering either SMD + PMR or only patients in PMR as responders. We found that only irRC response (PR + SD) was significantly associated with OS (hazard ratio, 0.293; 95% confidence interval, 0.121–0.709, $P = 0.004$). A borderline association was observed between PERCIST response and OS (hazard ratio, 0.355, 95% confidence interval, 0.103–1.224, $P = 0.066$).
No differences were highlighted in the classification of irRC PD by PERCIST and imPERCIST when only PMR patients were considered responders. Of note, among irRC PD, PMR patients showed a significantly longer OS ($P = 0.018$). In fact irRC PD patients had a median survival of 13.2 mo in PET responders versus 6.06 mo in PET nonresponders, suggesting that PET-based classification maintains some prognostic significance once irRC is adjusted for (**Figs. 6A–6C** and Supplemental Figs. 1 and 2; supplemental materials are available at http://jnm.snmjournals.org). Finally, no residual association with RECIST was observed ($P = 0.60$).

DISCUSSION

The present study aimed to provide a direct comparison between PET-based and CT-based response at the first evaluation in a group of patients with advanced NSCLC treated with nivolumab. Standard RECIST and PERCIST, as well as immunotherapy-modified irRC and imPERCIST, were used for CT- and PET-based response, respectively. A low overall concordance between CT-based and PET-based response was demonstrated, with the highest disagreement observed when PET-based criteria defined SMD or PMR. Of note, whereas SMD seems to be a rather uninformative label because it includes patients with largely variable OS, PMR was demonstrated to identify a subgroup of patients with longer OS irrespective of the results obtained on both CT-based criteria.
The low concordance between PET- and CT-based criteria was at least partially related to the reduced classification as PMD by PET criteria compared with RECIST and irRC. This

disagreement was markedly evident in cases classified as SMD or PMR by PERCIST or imPERCIST. Indeed, 47% and 52% of patients defined as SMD were defined as PD by both RECIST and irRC, and we found a lower cumulative survival for both of these subgroups of SMD patients. The lack of clinically prognostic relevance for SMD is further confirmed by the fact that the SMD and PMD curves were completely stackable in terms of OS, thus further underlining the false overstatement in terms of outcome. Indeed, SMD patients represent a highly heterogeneous group including both nonresponders, whose lesions did not significantly modify their metabolism because of lack of sensitivity to treatment, and responders in the earliest stages of response, who were not classified as PMR possibly because of immune cell infiltration resulting in a relative increase in ^{18}F-FDG uptake. The high heterogeneity in the time course of response to ICIs might further complicate the interpretation of PERCIST SMD at a single-patient level. In fact, a well-known potential feature of therapies targeting immune checkpoint pathways is cell infiltration, which may result in both the appearance of new lesions and an increase in lesion dimensions and thus may be confused with PD according to RECIST (**19**). Actually, the presence of new ^{18}F-FDG–avid lesions also results in progression when the same cases are assessed for response with ^{18}F-FDG PET by means of PERCIST, and in light of the analogy between the 2 response criteria, the inadequacy of RECIST in patients with new lesions was expected to be reflected also in PERCIST. In fact, in the present study, 8 patients were defined as PD according to both RECIST and PERCIST because of new CT-evident, markedly ^{18}F-FDG–avid lesions. To overcome this limitation, as previously introduced for CT-based criteria (**7**,**8**), in imPERCIST the appearance of new lesions alone did not result in PMD and new lesions were included in the sum of lean-body-mass–corrected peak SUV (SULpeak) if they showed higher uptake than in existing target lesions or if fewer than 5 target lesions were detected on the baseline scan. Similarly, Goldfarb et al. recently proposed immune PERCIST (**20**), accounting for unconfirmed and confirmed PD 4 wk after the initial ^{18}F-FDG–avid new-lesion finding. However, given the relatively low number of patients in the present study, only 2 patients were classified as SMD by imPERCIST among the subgroup who would have been classified as PMD by standard PERCIST. This slight difference prevents us from specifically investigating the predictive value of imPERCIST SMD versus PERCIST SMD. Accordingly, the evaluation of the added value of imPERCIST with respect to standard PERCIST deserves to be addressed in larger studies

In the present study, whereas SMD and PMD had a limited added value with respect to CT-based response, patients defined as PMR demonstrated a homogeneous better cumulative OS, thus suggesting that ^{18}F-FDG PET might open a new prognostic window in NSCLC patients treated with ICIs. A pooled analysis by Min et al. (**22**) demonstrated that significantly higher overall response rates are observed with PERCIST than with RECIST in NSCLC patients treated with chemotherapy. Although obtained in a new and different treatment setting, our results on PMR response are in keeping with Min's analysis, as we observed a 17.5% objective response rate with PERCIST and only 12% with both RECIST and irRC. Similarly, 15%–23% of patients considered PMR were defined as stable or even PD by RECIST and irRC, respectively, but all these patients were still alive at 12 mo. We further underline that in the present study, PMR was evident in patients classified as RECIST PD since the first response, thus supporting the role of ^{18}F-FDG PET as an early predictor of the efficacy of nivolumab in NSCLC patients (even before obtaining a further, 1-mo delayed, evaluation as required by irRC). Therefore, the present findings extend to NSCLC patients treated with nivolumab, as in the evidence provided by Sachpekidis et al., who reported that ^{18}F-FDG PET can be predictive of final treatment response in 18 of 22 patients with metastatic melanoma after 2 cycles of ipilimumab (**23**). It should be also noted that patients classified as CT-based PR (with both criteria) had the worst survival curves. However, the low number of

patients included in these categories prevent inference of more general comments on the trend of survival curves in patients with CT-based PR at first response.
Although larger groups of patients should be analyzed and a more comprehensive approach to ^{18}F-FDG PET might include the evaluation of other PET-based variables, our results suggest the potential capability of ^{18}F-FDG PET to reduce the risk of misclassifying pseudoprogression as PD. In fact, in our post hoc analysis, both PET-based criteria were able to identify a subgroup of patients with longer OS classified as irRC PD, thus suggesting that PET-based classification might maintain some prognostic significance once irRC is adjusted for. In previous studies on the use of PET-based response in melanoma patients treated with ICIs, as many as half the patients showing residual disease on CT had negative findings on ^{18}F-FDG PET (**23**–**26**). Indeed, the risk for PD misclassification after ICIs was higher in initial observations, which found that 10% of melanoma patients who, during ipilimumab, would have been misclassified as PD by World Health Organization criteria showed a clinical response (including PR and SD) (**27**). Subsequently, the rate of pseudoprogression was reported as 0%–6% in NSCLC treated with ICIs, and at present, pseudoprogression should be considered mainly when the clinical condition of a patient in apparent radiologic progression is concomitantly improving (**28**,**29**). In keeping with this evidence, recent ^{18}F-FDG PET studies introduced evaluation of clinical benefit into the definition of response (**25**,**26**,**30**). In this framework, our findings do not support the routine use of ^{18}F-FDG PET to monitor response to nivolumab in all NSCLC patients but do support the role of ^{18}F-FDG PET in the prognostic stratification of patients defined as having PD according to CT-based criteria. Accordingly, ^{18}F-FDG PET evaluation might play a role in patients with suspected pseudoprogression, although only the availability of a baseline ^{18}F-FDG PET scan may allow confirmation or exclusion of progression by means of ^{18}F-FDG PET later in the disease course.

The present study had some drawbacks. The first was the relatively small population. However, the monocentric nature of the study represents one of its strengths. Accordingly, all 48 enrolled patients underwent ^{18}F-FDG PET using the same PET/CT scanner, avoiding the possible influence of interscanner variability on ^{18}F-FDG PET results. Second, the present preliminary findings apply to the PERCIST and imPERCIST-based assessment of ^{18}F-FDG PET and may not necessary be applied to other methods used to evaluate PET-based response to ICIs. In fact, even in NSCLC patients treated with conventional chemotherapy, no consensus has yet been reached on the best PET-based approach to assess response to therapy, and methods based on other criteria should also be considered, such as metabolic tumor volume assessment or the criteria of the European Organisation for Research and Treatment of Cancer (**30**,**31**). Indeed, Kaira et al. (**10**) addressed the role of several PET-derived parameters, including metabolic tumor volume and total lesion glycolysis, and highlighted the value of total lesion glycolysis as an independent predictor of PFS and OS in patients with NSCLC treated with nivolumab.

However, none of newly proposed criteria or PET-based variables have yet been validated, and in a recent position paper by Aide et al., the use of either PERCIST or the PET response criteria of the European Organisation for Research and Treatment of Cancer is still suggested for evaluating ^{18}F-FDG uptake changes in target lesions in patients treated with ICIs (**32**). Moreover, in the recommendations by Aide et al., the addition of computation of metabolic tumor volume and total lesion glycolysis to the evaluation of PET response is mentioned just as a possibility. Therefore, given that the best method to compute metabolic tumor volume even in patients treated with conventional chemotherapy is still a matter of debate, in the present study we decided to explore several different CT and PET approaches based on

better-established measurements (i.e., the lean-body-mass–corrected peak SUV). With respect to CT-based response evaluation, besides irRC, the use of immune-RECIST for CT-response evaluation in patients treated with ICIs has more recently been proposed. Actually, irRC was among the first set of criteria proposed, and at the time that the design of the present study was defined and approved, there were no other accepted or validated response criteria. The use of irRC was therefore part of the prospective design of the study, thus obviously influencing the recruited patients' management. The added value of PET-based criteria with respect to immune-RECIST was thus not addressed in the present study. Another potential limitation of the present study relates to the multivariate analysis. In fact, we did not consider other potential prognostic factors in the cohort to be capable of influencing OS or, consequently, the impact of the different response criteria investigated. Finally, one of the binary classifications that was applied in the post hoc analysis was based on the inspection of the Kaplan–Meier curves showing that patients classified as SD by RECIST and irRC had a survival experience like that of patients in PR, whereas those classified as SMD by PERCIST resembled patients in PMD. Therefore, we evaluated the residual prognostic role of PET in irRC response categories considering either SMD + PMR or only patients in PMR as responders. The results of the Cox analysis in the latter patients' subgrouping should thus be considered with caution because of the risk of overfitting, since the model was derived from the data.

CONCLUSION

Even though the present study did not support the routine use of ^{18}F-FDG PET in the general population of NSCLC patients treated with nivolumab, it supports the prognostic potential of the metabolic response assessment and the importance of acquiring baseline ^{18}F-FDG PET data to compare with posttreatment examinations in complex cases for which therapeutic decision making can potentially be improved.

DISCLOSURE

Carlo Genova receives honoraria from Astra Zeneca, BMS, Boehringer Ingelheim, MSD, and Roche. Francesco Grossi has an advisory role on ad hoc advisory boards or as a consultant (last 3 y) for Eli Lilly, Roche, Boehringer Ingelheim, AstraZeneca, Pierre Fabre, BMS, MSD, and Novartis Francesco Grossi also receives honoraria from giving seminars or talks to industry (last 3 y) from Eli Lilly, Roche, Boehringer Ingelheim, AstraZeneca, Pierre Fabre, AMGEN, Celgene, BMS, and MSD. Silvia Morbelli received speaker honoraria from GE Healthcare (2017). No other potential conflict of interest relevant to this article was reported.

KEY POINTS

- **QUESTION:** What is the role of ^{18}F-FDG PET in the evaluation of response to therapy in NSCLC patients treated with ICIs?

- **PERTINENT FINDINGS:** The present single-institution translational research trial strongly suggests that metabolic response assessment has prognostic potential. Indeed, PMR assessed by PERCIST predicted longer OS than did CT-based PR.

- **IMPLICATIONS FOR PATIENT CARE:** It might be useful to acquire baseline ^{18}F-FDG PET data on NSCLC patients who are candidates for ICIs for comparison with posttreatment examinations in complex cases for which therapeutic decision making can potentially be improved.

REFERENCES

1. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45: 228–247.

2. Hotta K, Kato Y, Leighl N, et al. Magnitude of the benefit of progression-free survival as a potential surrogate marker in phase 3 trials assessing targeted agents in molecularly selected patients with advanced non-small cell lung cancer: systematic review. PLoS One. 2015;10:e0121211.

3. Akamatsu H, Mori K, Naito T, et al. Progression-free survival at 2 years is a reliable surrogate marker for the 5-year survival rate in patients with locally advanced non-small cell lung cancer treated with chemoradiotherapy. BMC Cancer. 2014;14:18.

4. Nishino M, Giobbie-Hurder A, Gargano M, et al. Developing a common language for tumor response to immunotherapy: immune-related response criteria using unidimensional measurements. Clin Cancer Res. 2013;19:3936–3943.

5. Beaver JA, Hazarika M, Mulkey F, et al. Patients with melanoma treated with an anti-PD-1 antibody beyond RECIST progression: a US Food and Drug Administration pooled analysis. Lancet Oncol. 2018;19:229–239.

6. Tazdait M, Mezquita L, Lahmar J, et al. Patterns of responses in metastatic NSCLC during PD-1 or PDL-1 inhibitor therapy: comparison of RECIST 1.1, irRECIST and iRECIST criteria. Eur J Cancer. 2018;88:38–47.

7. Wolchok JD, Hoos A, O'Day S, et al. Guidelines for the evaluation of immune therapy activity in solid tumors: immune-related response criteria. Clin Cancer Res. 2009;15:7412–7420.

8. Seymour L, Bogaerts J, Perrone A, et al. iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. Lancet Oncol. 2017;18: e143–e152.

9. Bauckneht M, Piva R, Sambuceti G, Grossi F, Morbelli S. Evaluation of response to immune checkpoint inhibitors: is there a role for positron emission tomography? World J Radiol. 2017;9:27–33.

10. Kaira K, Higuchi T, Naruse I, et al. Metabolic activity by 18F-FDG-PET/CT is predictive of early response after nivolumab in previously treated NSCLC. Eur J Nucl Med Mol Imaging. 2018;45:56–66.

11. Eshghi N, Lundeen TF, Kuo PH. Dynamic adaptation of tumor immune response with nivolumab demonstrated by 18F-FDG PET/CT. Clin Nucl Med. 2018;43: 114–116.

12. Ito K, Teng R, Sch¨oder H, et al. 18F-FDG PET/CT for monitoring of ipilimumab therapy in patients with metastatic melanoma. J Nucl Med. 2019;60:335–341.

13. Higuchi M, Owada Y, Inoue T, et al. FDG-PET in the evaluation of response to nivolumab in recurrent non-small-cell lung cancer. World J Surg Oncol. 2016;14: 238.

14. Brahmer J, Reckamp KL, Baas P, et al. Nivolumab versus docetaxel in advanced squamous-cell non-small-cell lung cancer. N Engl J Med. 2015;373:123–135.

15. Wahl RL, Jacene H, Kasamon Y, Lodge MA. From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors. J Nucl Med. 2009;50(suppl 1):122S–150S.

16. Boellaard R, Delgado-Bolton R, Oyen WJ, et al. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. Eur J Nucl Med Mol Imaging. 2015;42:328–354.

17. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159–174.

18. Bergeri I, Michel R, Boutin JP. Everything (or almost everything) about the kappa coefficient [in French]. Med Trop Rev. Med Trop (Mars). 2002;62:634–636.

19. Lara PN Jr, Redman MW, Kelly K, et al.; Southwest Oncology Group. Disease control rate at 8 weeks predicts clinical benefit in advanced non-small-cell lung cancer: results from Southwest Oncology Group randomized trials. J Clin Oncol. 2008;26:463–467.

20. Borcoman E, Kanjanapan Y, Champiat S, et al. Novel patterns of response under immunotherapy. Ann Oncol. 2019;30:385–396.

21. Goldfarb L, Duchemann B, Chouahnia K, Zelek L, Soussan M. Monitoring antiPD-1-based immunotherapy in non-small cell lung cancer with FDG PET: introduction of iPERCIST. EJNMMI Res. 2019;9:8.

22. Min SJ, Jang HJ, Kim JH. Comparison of the RECIST and PERCIST criteria in solid tumors: a pooled analysis and review. Oncotarget. 2016;7:27848–27854.

23. Sachpekidis C, Anwar H, Winkler J, et al. The role of interim 18F-FDG PET/CT in prediction of response to ipilimumab treatment in metastatic melanoma. Eur J Nucl Med Mol Imaging. 2018;45:1289–1296.

24. Sachpekidis C, Kopp-Schneider A, Hakim-Meibodi L, DimitrakopoulouStrauss A, Hassel JC. 18F-FDG PET/CT longitudinal studies in patients with advanced metastatic melanoma for response evaluation of combination treatment with vemurafenib and ipilimumab. Melanoma Res. 2019;29:178– 186.

25. Ito K, Teng R, Sch ¨oder H, et al. 18F-FDG PET/CT for monitoring of ipilimumab therapy in patients with metastatic melanoma. J Nucl Med. 2019;60:335– 341.

26. Cho SY, Lipson EJ, Im HJ, et al. Prediction of response to immune checkpoint inhibitor therapy using early-time-point 18F-FDG PET/CT imaging in patients with advanced melanoma. J Nucl Med. 2017;58:1421–1428.

27. Sachpekidis C, Hassel JC, Dimitrakopoulou-Strauss A. 18F-FDG PET/CT reveals disease remission in a patient with ipilimumab-refractory advanced melanoma treated with pembrolizumab. Clin Nucl Med. 2016;41:156–158.

28. Wolchok JD, Hoos A, O'Day S, et al. Guidelines for the evaluation of immune therapy activity in solid tumors: immunerelated response criteria. Clin Cancer Res. 2009;15:7412–7420.

29. Vrankar M, Unk M. Immune RECIST criteria and symptomatic pseudoprogression in non-small cell lung cancer patients treated with immunotherapy. Radiol Oncol. 2018;52:365–369.

30. Castello A, Grizzi F, Qehajaj D, Rahal D, Lutman F, Lopci E. 18F-FDG PET/CT for response assessment in Hodgkin lymphoma undergoing immunotherapy with checkpoint inhibitors. Leuk Lymphoma. 2019;60:367–375.

31. Anwar H, Sachpekidis C, Winkler J, et al. Absolute number of new lesions on 18F-FDG PET/CT is more predictive of clinical response than SUV changes in metastatic melanoma patients receiving ipilimumab. Eur J Nucl Med Mol Imaging. 2018;45:376–383.

32. Aide N, Hicks RJ, Le Tourneau C, Lheureux S, Fanti S, Lopci E. FDG PET/CT for assessing tumour response to immunotherapy: report on the EANM symposium on immune modulation and recent review of the literature. Eur J Nucl Med Mol Imaging. 2019;46:238–250.

TABLE 1
Patient Characteristics

Characteristic	Parameter	Data
Age (y)		**70 (range, 44–85)**
Sex	**Female**	**15/48 (31%)**
	Male	**33/48 (69%)**
ECOG	**0**	**20/48 (42%)**
	1	**25/48 (52%)**
	2	**3/48 (6%)**
Smoking status	**Never smoker**	**7/48 (15%)**
	Former smoker	**24/48 (50%)**
	Smoker	**17/48 (35%)**
Histology	**Squamous**	**11/48 (23%)**
	Nonsquamous	**37/48 (77%)**
Prior lines of therapy	**1**	**17/48 (35%)**
	2	**13/48 (27%)**
	≥3	**18/48 (38%)**

ECOG = Eastern Cooperative Oncology Group performance status.

TABLE 2
Concordance Between RECIST and PERCIST Scores at First Evaluation

	PERCIST			
RECIST	**PMD**	**SMD**	**PMR**	**Total**
PD	14 (29.2%)	11 (22.9%)	2 (4.2%)	27 (56.3%)
SD	1 (2.1%)	11 (22.9%)	5 (10.4%)	17 (35.4%)
PR	0 (0%)	1 (2.1%)	3 (6.3%)	4 (8.3%)
Total	15 (31.3%)	23 (47.9%)	10 (20.8%)	48 (100%)

TABLE 3
Concordance Between RECIST and imPERCIST Scores at First Evaluation

	imPERCIST			
RECIST	**PMD**	**SMD**	**PMR**	**Total**
PD	**13 (27.1%)**	**12 (25%)**	**2 (4.2%)**	**27 (56.3%)**
SD	**0 (0%)**	**12 (25%)**	**5 (10.4%)**	**17 (35.4%)**
PR	**0 (0%)**	**1 (2.1%)**	**3 (6.3%)**	**4 (8.3%)**
Total	**13 (27.1%)**	**25 (52.1%)**	**10 (20.8%)**	**48 (100%)**

TABLE 4
Concordance Between irRC and PERCIST Scores at First Evaluation

	PERCIST			
irRC	**PMD**	**SMD**	**PMR**	**Total**
PD	**10 (20.8%)**	**12 (25%)**	**2 (4.2%)**	**24 (50%)**
SD	**5 (10.4%)**	**10 (20.8%)**	**6 (12.5%)**	**21 (43.8%)**
PR	**0 (0%)**	**1 (2.1%)**	**2 (4.2%)**	**3 (6.3%)**
Total	**15 (31.3%)**	**23 (47.9%)**	**10 (20.8%)**	**48 (100%)**

TABLE 5
Concordance Between irRC and imPERCIST Scores at First Evaluation

	imPERCIST			
irRC	**PMD**	**SMD**	**PMR**	**Total**
PD	**10 (20.8%)**	**12 (25%)**	**2 (4.2%)**	**24 (50%)**
SD	**3 (6.2%)**	**12 (25%)**	**6 (12.5%)**	**21 (43.8%)**
PR	**0 (0%)**	**1 (2.1%)**	**2 (4.2%)**	**3 (6.3%)**
Total	**13 (27.1%)**	**25 (52.1%)**	**10 (20.8%)**	**48 (100%)**

FIGURE 1.
Schematic of study design. The first evaluation was performed with CT and ^{18}F-FDG PET after 8 wk (4 cycles of nivolumab). If patient experienced PD according to CT criteria, evaluation was repeated after a further 2 cycles (4 wk). Otherwise, in cases of SD, PR, or CR, evaluation was repeated every 4 cycles.

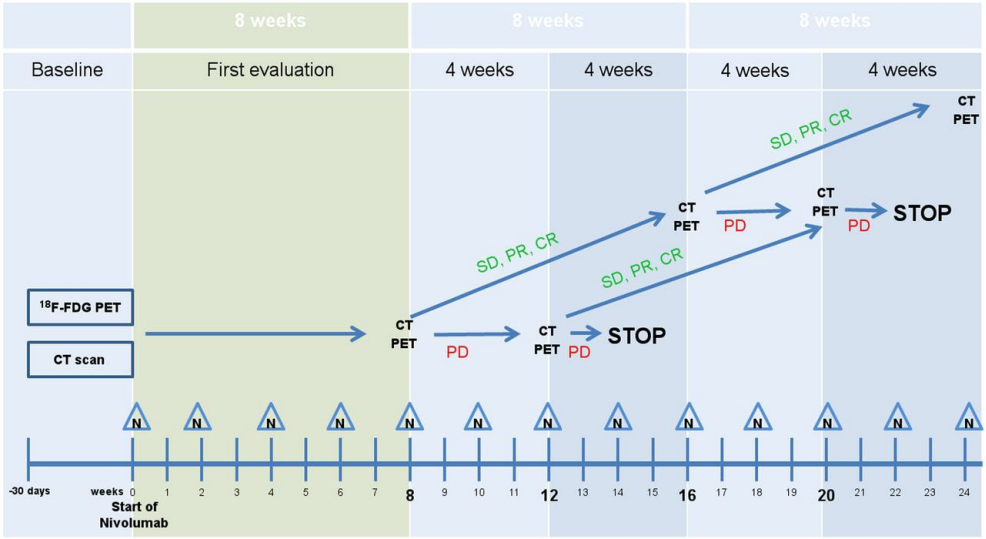


FIGURE 2.
Patients eligible for study.

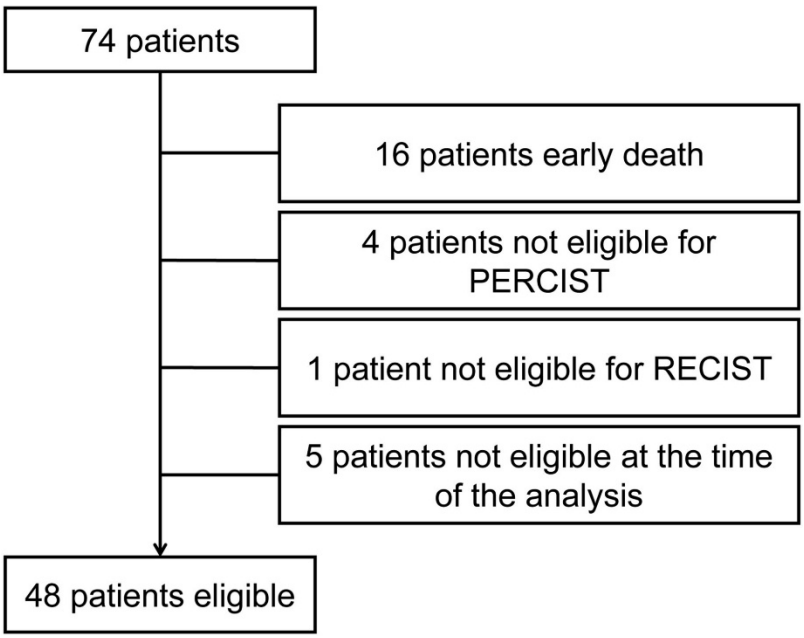


FIGURE 3.

Representative PET/CT images of 3 patients showing discordant response at CT-based and PET-based evaluation or between PERCIST and imPERCIST. Patient 1 was classified as PD with RECIST because of significant increase in dimensions of perihilar lesion in right lung; however, marked reduction of lean-body-mass–corrected peak SUV allowed us to classify this patient as PMR according to both PERCIST and imPERCIST. In contrast, patient 2 showed marked lesion shrinkage, but lesion metabolism was substantially stable (possibly from inflammatory infiltration) or even mildly increased in the case of right pleural metastasis. This patient was classified as SMD with both PERCIST and imPERCIST, but response was more evident on CT images. Patient 3 was differently classified by PERCIST and imPERCIST. At first response (2 mo after therapy), he was classified as SMD by imPERCIST but, because of new lesions, was considered PERCIST PMD. After 1 mo more, new lesions disappeared and lesions present since baseline showed significant reduction in both dimensions and metabolism.

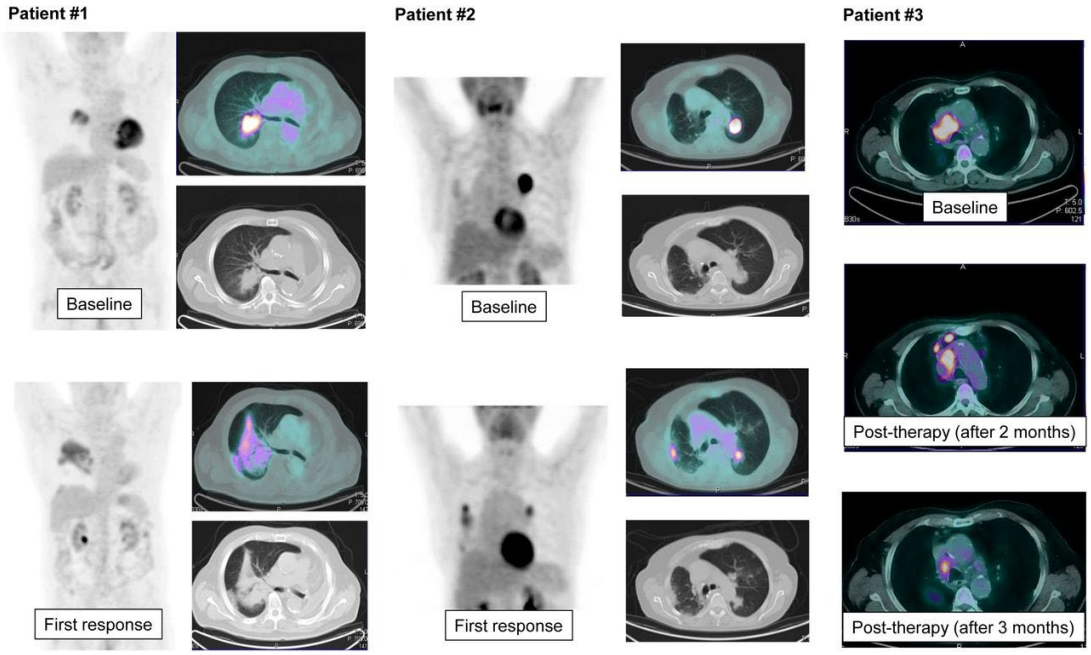


FIGURE 4.
OS curves of overall population of study.

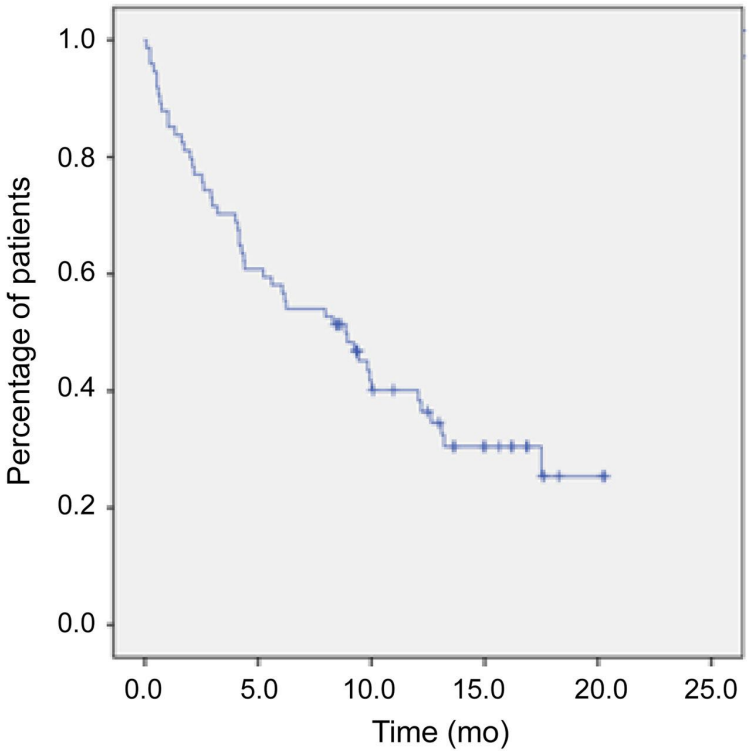


FIGURE 5.
OS according to CT-based and PET-based criteria at first evaluation: RECIST (A), irRC (B), PERCIST (C), and imPERCIST (D).

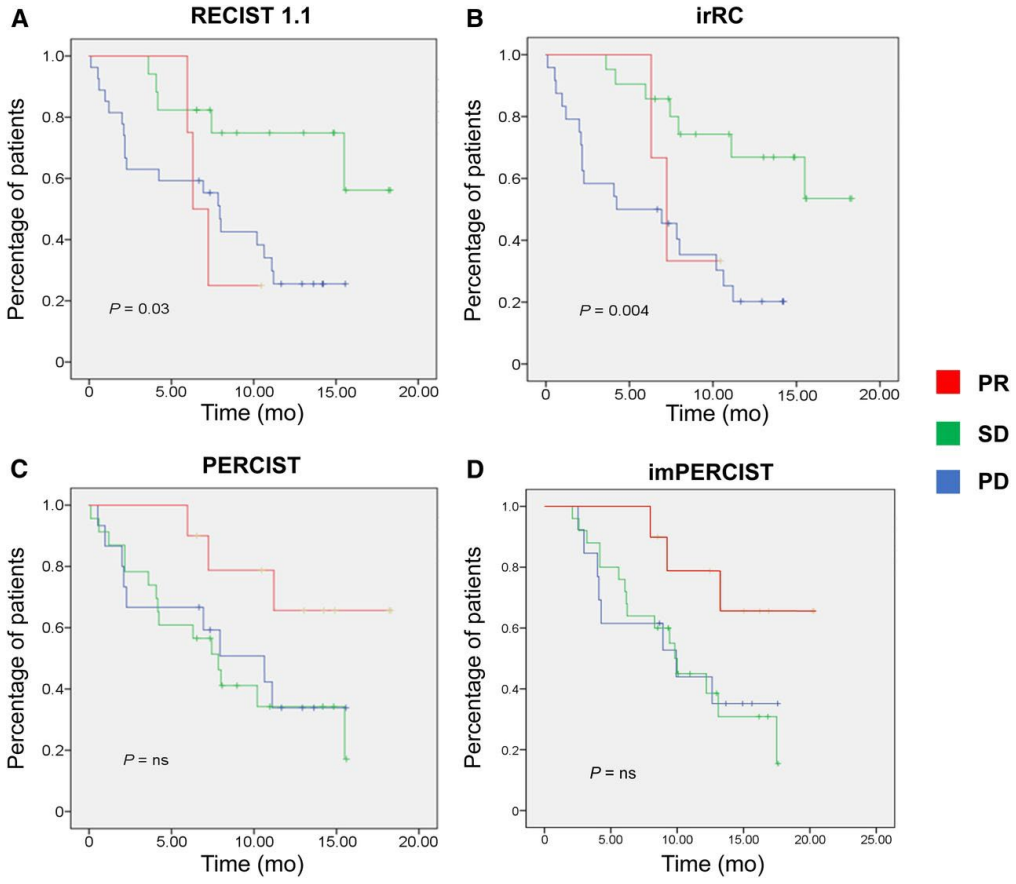
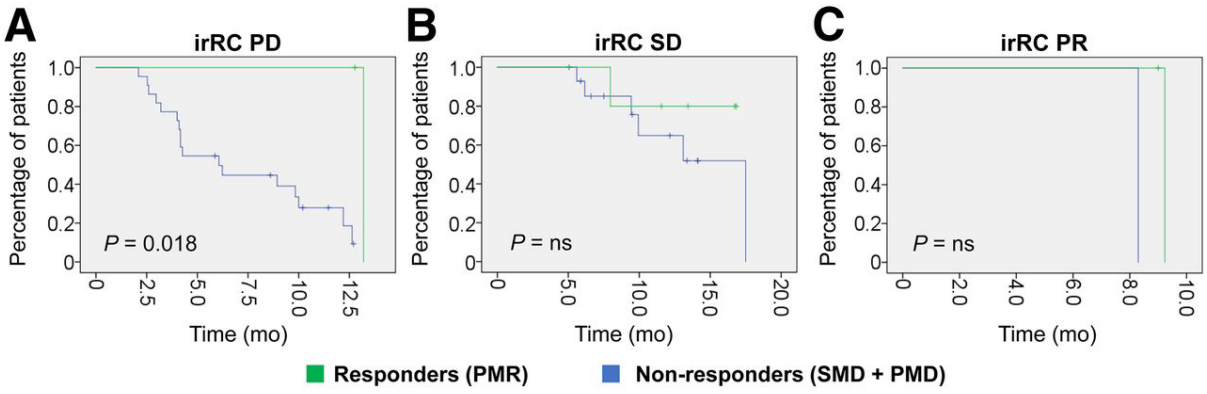


FIGURE 6.
Survival OS curves assessed by means of binary classification of first response according to both PET-based criteria in responders (PMR) or nonresponders (SMD + PMD), in patients classified according to irRC PD (A), irRC SD (B), and irRC PR (C).


A
irRC PD
Percentage of patients
1.0
0.8
0.6
0.4
0.2
0
P = 0.018
0
2.5
5.0
7.5
10.0
12.5
Time (mo)
B
irRC SD
Percentage of patients
1.0
0.8
0.6
0.4
0.2
0
P = ns
0
5.0
10.0
15.0
20.0
Time (mo)
C
irRC PR
Percentage of patients
1.0
0.8
0.6
0.4
0.2
0
P = ns
0
2.0
4.0
6.0
8.0
10.0
Time (mo)
Responders (PMR)
Non-responders (SMD + PMD)