Journal of Nuclear Medicine, published on December 5, 2019 as doi:10.2967/jnumed.119.233056

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

Running title: PET vs CT after immunotherapy for NSCLC

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], Francesco Grossi[4]

1. Lung Cancer Unit, IRCCS Policlinico San Martino, Genoa, Italy

2. Department of Clinical, surgery and experimental sciences, division of experimental Pathology and Oncology, University of 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, Italy

8. Nuclear Medicine Unit, Humanitas Clinical and Research Center - IRCCS, Rozzano, Italy

Total word count: 5051

First author:

Giovanni Rossi, MD
Medical Oncology resident
Lung Cancer Unit
IRCCS Policlinico San Martino, Genoa, Italy
Phone: +390105553038
e-mail: giovanni.rossi.1689@gmail.com

Corresponding author:

Prof. Silvia Morbelli, MD, PhD
Nuclear Medicine Unit
Department of health sciences (DISSAL), University of Genoa
IRCCS Policlinico San Martino, Genoa, Italy
Phone: +390105553038

e-mail: silvia.morbelli@unige.it

Conflicts of interest: CG: Honoraria from Astra Zeneca, BMS, Boehringer Ingelheim, MSD, Roche; FG: Advisory Role: Ad Hoc Advisory Boards/Consultations (last 3 years), Eli Lilly, Roche, Boehringer Ingelheim, AstraZeneca, Pierre Fabre, BMS, MSD, Novartis Honoraria: Seminar/Talks to Industry (last 3 years), Eli Lilly, Roche, Boehringer Ingelheim, AstraZeneca, Pierre Fabre, AMGEN, Celgene, BMS, MSD, Research Funding (last 3 years); SM: Speaker honoraria from General Electric Healthcare (2017). No other potential conflicts of interest relevant to this article exist.

ABSTRACT

Due to their peculiar mechanism of action, the evaluation of radiological response to immune checkpoint inhibitors (ICI) presents many challenges in solid tumors. We aimed to compare the evaluation of first response to Nivolumab by means of CT-based criteria with respect to fluorodeoxyglucose positron emission tomography (FDG-PET) response criteria in non-small-cell lung cancer (NSCLC) patients.

Methods. 72 patients with advanced NSCLC were recruited in a mono-institutional ancillary trial within the expanded access program (EAP; NCT02475382) for Nivolumab. Patients underwent CT scan and FDG-PET at baseline and after 4 cycles (first evaluation). In case of progressive disease (PD), an additional evaluation was performed after two further cycles in order to confirm progression. We evaluated the response to treatment with CT scan by means of response evaluation criteria in solid tumors (RECIST) 1.1 and Immuno-related Response Criteria (IrRC) and with FDG-PET by means of PERCIST and immunotherapy-modified-PERCIST (imPERCIST) criteria. The concordance between CT- and PET-based criteria and the capability of each method to predict overall survival (OS) were evaluated.

Results. 48/72 patients were evaluable for first response assessment with both PET- and CT-based criteria. We observed low concordance between CT- and PET-based criteria (Kappa value of 0.346 and 0.355 and Kappa value of 0.128 and 0.198 between PERCIST and imPERCIST versus RECIST and irRC respectively). Looking at OS, IrRC were more reliable to distinguish responders from non-responders. However thanks to the prognostic value of partial metabolic response assessed by both PERCIST and Immuno-PERCIST, PET-based response maintained prognostic significant in patients classified as progressive disease on the basis of irRC**.**

Conclusion. Even though the present study did not support the routine use of FDG-PET in the general population of NSCLC patients treated with ICI, it suggests the added prognostic value of the metabolic response assessment, potentially improving the therapeutic decision-making.

KEY WORDS: NSCLC; checkpoint inhibitors; Positron Emission Tomography; Computed Tomography.

INTRODUCTION

Response to therapy in solid tumors is conventionally monitored by means of morphological imaging. Traditionally, shrinkage in tumor burden describes treatment success and if not seen, patients are assumed to be non-responders. The two most widely used systems for the classification of tumor shrinkage, proposed by the World Health Organization and the Response Evaluation Criteria in Solid Tumors (RECIST), were developed to standardize response evaluation in phase II clinical trials (1). However, while patients who respond to treatment are known to have a better prognosis, the validity of objective response to chemotherapy as a surrogate endpoint of survival is more contradictory especially in some clinical settings (2).

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

In this scenario, while the morphology-based criteria are in course of further validation, less evidence is available on the added value of 18F-Fluorodeoxyglucose Positron Emission Tomography (FDG-PET) in patients treated with ICI (9). Despite the potential occurrence of inflammatory infiltration and related tumor changes that might also hamper reliability of PET signal, FDG-PET might be able to capture response to ICI in specific subgroups of patients. However, the exact positioning of FDG-PET in the flowchart of patients treated with ICI and its cost-effectiveness with respect to conventional morphological assessment still need to be defined. Moreover, PET Response Criteria in Solid Tumors (PERCIST) were introduced in 2009 as guidelines for structured assessment of response to therapy with 18F-FDG PET in oncology but the frequency and impact of pseudoprogression as seen by PERCIST in patients treated with ICI is not well documented. Similarly, immunotherapy-modified PERCIST (imPERCIST) criteria have been proposed but 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 NSCLC patients treated with ICI (10-13). The aim of the present study was to assess, in patients with advanced NSCLC, the frequency and pattern of FDG-PET-based first response to Nivolumab, a fully human IgG4 program-death-1 (PD-1) antibody (14), and to compare them with both RECIST 1.1 and irRC CT-based

criteria. To this aim, PET-response was both evaluated with standard PERCIST and imPERCIST criteria (15). As a secondary aim the correlation between PET- or CT-based criteria and patients' 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 mono-institutional study conducted within the expanded access program (EAP) for Nivolumab (NCT02475382). According the specific study design was approved by the research committee of Regione Liguria (Italy) and only patients enrolled at our institution were included in the study. All the enrolled patients gave a written informed consent to participate in this study. Nivolumab was provided by Bristol-Meyer-Squibb within the EAP in NSCLC. Major inclusion criteria were: age ≥18 years, histologically or cytologically confirmed NSCLC, clinical stage IIIb or IV (according to TNM v7.0), at least one previous line of therapy, at least one measurable lesion by RECIST 1.1, previously treated or stable brain metastasis from at least 2 weeks before the treatment with Nivolumab that did not need steroids with more than 10 mg/die of prednisone or equivalent. Exclusion criteria were Eastern Cooperative Oncology Group (ECOG) scale of performance status ≥3, meningeal carcinomatosis, active autoimmune disease or syndrome that needed daily steroids treatment (excepted for diabetes mellitus type I, hypothyroidism only requirement hormone replacement), previous line of therapy with ICI, and finally the administration of a life attenuated vaccine within the 30 days before the first Nivolumab administration.

CT and FDG-PET were performed as detailed in (16) and in the Supplementary Materials within 30 days before starting therapy with Nivolumab 3 mg/kg every 14 days. Imaging was repeated after 4 cycles (first evaluation) and then every 4 cycles. Response to treatment was evaluated by CT using the RECIST 1.1 (3) (from now on defined simply as RECIST) and IrRC (7), and by FDG-PET using PERCIST (15), as detailed in the Supplementary Materials. If patients experienced progressive disease (PD) by RECIST, the protocol required to repeat CT and FDG-PET after two additional cycles to confirm PD (IrRC) and to treat patients beyond progression, in case of clinical benefit. Patients with stable disease (SD) or Partial Response (PR) repeated CT and FDG-PET every 4 cycles (a flow-chart showing a schematic representation of the study design is reported in Figure 1). Only patients with at least one post-therapy evaluation with both CT and 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 post-therapy evaluation (either CT or FDG-PET) was not available.

Definition of group response

CT were interpreted by physicians experienced in response evaluation with both RECIST and irRC (GR, FG, GC), blinded to PET/CT results. Similarly, FDG-PET were interpreted according to standard PERCIST and imPERCIST criteria by two nuclear medicine physicians experienced in PERCIST-based response evaluation (MB, SM) blinded to CT results. Response criteria are detailed elsewhere (3,7,15) and summarized in Supplementary Materials.

Statistical analysis

Patients without, at least, one CT and one FDG-PET 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 first evaluation by means of FDG-PET and were excluded. Accordingly, survival curves start at 2 months after the first treatment administration. The concordance between CT-based criteria and PERCIST was investigated computing a Cohen's K coefficient. Unweighted kappa values are reported using the benchmarks of Landis and Koch (17, 18). Concordance between the first evaluation by RECIST 1.1 and by IrRC versus PERCIST and imPERCIST were calculated. In order to understand the prognostic value of treatment response assessed by FDG-PET, alone or in addition to CT imaging, univariate OS curves by PERCIST response were computed according to Kaplan Meier and compared with the Log Rank test. To evaluate the relative contribution of each of the three classification systems to prognosis, a multivariate Cox Proportional Hazard 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, also the possibility of a synergy between the imaging and metabolic techniques in improving the prognostic ability of the model were evaluated by including in the model the appropriate interaction terms and by evaluating the resulting modification of its likelihood. All tests are two sided. Analyses were conducted with IBM - SPSS release 23.

RESULTS

Out of 74 patients included in the EAP, 48 underwent both CT and FDG-PET at baseline, during their treatment course and were characterized by lesions characterized by metabolism suitable for evaluation by means of PERCIST. Therefore only 48 patients were included in the present analysis (Figure 2). Main patients' characteristics are summarized in Table 1. At the first CT evaluation, 27/48 (56%) of patients were classified as PD, 17 (36%) as SD and 4 (8%) experienced 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 Complete Metabolic Response (CMR), 11 (22%) as Partial Metabolic Response (PMR), 22 (46%) with Stable Metabolic Disease (SMD) and 14 (29%) as Progressive Metabolic Disease (PMD). Classification according to imPERCIST was substantially overlapping PERCIST response with only two patients considered PMD according

to PERCIST and conversely classified as SMD according to imPERCIST (imPERCIST classification in the entire cohort: 1 patient (2%) was classified as CMR, 11 (22%) as PMR, 24 (50%) with SMD and 12 (25%) as PMD).

Overall, the agreement between first response evaluated by RECIST and both PET-based criteria was moderate, as only 58% of the patients were similarly classified by the two evaluations (Kappa value of 0.346 with respect to PERCIST and 0.355 with respect to imPERCIST; p=0.001, Tables 2a and 3a). Out 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, out 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; Table 2b and 3b). As predictable, the two patients classified as SMD rather than PMD according to imPERCIST were also classified as SD according to IRrC. Two representative example showing 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 understand which method could be more reliable in predicting the prognosis, and if any improvement could be achieved by combining them. Mean OS within the whole population of the study are represented in Figure 4. On a per-criterion analysis, OS appeared to be similarly assessed by CT-based and PERCIST-based response criteria (Figure 3a-c). Patients classified as PD according to RECIST, IrRC, PERCIST and imPERCIST showed a uniformly poor prognosis, with a median OS of 8.9 months, 8.4 months, 9.3 and 9.9 months, respectively. However, differences were noticed among patients classified as SD. Indeed, in RECIST- and IrRC-SD patients, OS was similar or even better than that of patients achieving a PR. By contrast, in SMD, OS closely resembled that of patients with PMD (regardless of the used PET-based approach). Finally, longer OS is predicted by PERCIST and imPERCIST-PMR with respect to the few patients classified as RECIST- and IrRC-PR (Figure 5a-c).

Post-hoc analysis

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 dependent variable and RECIST, IrRC or PERCIST classes. Due to the small numbers involved, a binary classification of response was used, and patients were divided in “responders” or “non-responders” for each of the three criteria. In the framework of clinical trials for advanced NSCLC, CT-based disease control rate (DCR) at first response (CR + PR + SD) has demonstrate a significant positive predictive value (19). In the era of biologic “targeted” therapies and ICI, it has been further demonstrated that DCR metric more closely mirrors treatment effect than 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 ICI 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 of SMD whose curves demonstrated to largely overlap PMD curves both for PERCIST and imPERCIST criteria. Accordingly to fully exploit, 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. As a matter of fact, only IrRC response (PR+SD) was significantly associated with OS (HR=0.293, 95% CI: 0.121 to 0.709, p= 0.004) while a borderline association was observed between PERCIST response and OS (HR=0.355, 95% CI: 0.103 to 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 as 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 months in PET-responders vs 6.06 months in PET non-responders, suggesting that PET-based classification maintains some prognostic significance once irRC is adjusted for (Figure 6a-c and Supplemental Figure 1 and 2). 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 in a group of patients with advanced NSCLC treated with Nivolumab at first evaluation. Standard RECIST and PERCIST criteria as well as immunotherapy-modified irRC and imPERCIST criteria were used for CT and PET-based response respectively. 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 response. Of note, while SMD seems to be a rather uninformative label as it includes patients with largely variable OS, PMR demonstrated to identify a subgroup of patients with longer OS irrespective of the results obtained on both CT-based response criteria.

It should be noted that the low concordance between PET- and CT-based criteria was at least partially related to the reduced classification as PMD by PET criteria compared to 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 criteria and we found a lower cumulative survival for both these subgroups of SMD patients. The lack of clinical prognostic relevance of SMD is further confirmed by the fact that SMD and PMD curves are 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 non-responders, whose lesions did not significantly modify their metabolism due to lack of sensitivity to treatment, and responders in earliest stages of response who are not classified as PMR possibly due

to immune cell infiltration resulting in a relative increase in FDG uptake. The high heterogeneity in the time-course of response to ICI 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 the cell infiltration which may result in both appearance of new lesions or increase in lesions' dimensions and thus may be confused with progressive disease according to RECIST (19). Actually, the presence of new FDG-avid lesions results in progression also when the same cases are assessed for response with FDG-PET by means of PERCIST criteria, and in the light of the analogy between the two response criteria, the inadequacy of RECIST criteria in patients with new lesions was expected to be reflected also in PERCIST evaluation. In fact, in the present study, 8 patients were defined as progressive disease according to both RECIST and PERCIST criteria due to new CT evident, markedly uptaking FDG-avid lesions. To overcome this limitation, as previously introduced for CT-based criteria (7,8), in the imPERCIST criteria the appearance of new lesions alone did not result in PMD and new lesions were included in the sum of SULpeak if they showed higher uptake than existing target lesions or if fewer than 5 target lesions were detected on the baseline scan. Similarly, Goldfarb et al. recently proposed the iPERCIST criteria (20), accounting for unconfirmed and confirmed progressive disease after 4 weeks from the initial FDG-avid new-lesion finding. It should be noted, however, that, given the relatively low number of patients, in the present study only two patients were classified as SMD by imPERCIST among the subgroup that would have been classified as PMD by standard PERCIST criteria. This slight difference prevents us from a specific investigation of the predictive value of SMD-imPERCIST versus SMD-PERCIST. Accordingly, the evaluation of the added value of imPERCIST with respect to standard PERCIST approach deserve to be addressed in larger studies

In the present study, while SMD and PMD demonstrated to have a limited added value with respect to CT-based response, patients defined as in PMR demonstrated a homogeneous better cumulative OS thus suggesting that FDG-PET might open a new prognostic window in NSCLC patients treated with ICI. A recent pooled analysis by Min et al. (21) 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 17.5% of objective response rate with PERCIST and only 12% with both RECIST and irRC. Similarly, from 15 to 23% of patients considered in PMR were defined as stable or even PD by RECIST and irRC, respectively, but all these patients were still alive at 12 months. It is worthwhile to further underline that in the present study, PMR was evident in patients classified as RECIST-PD since first response thus supporting the role of FDG-PET as an early predictor of the efficacy of Nivolumab in NSCLC patients (even before obtaining a further, one month-delayed evaluation as required by irRC criteria). Therefore, the present findings extend to NSCLC patients treated with Nivolumab the evidence provided by Sachpekidis et al. who reported that FDG-PET can be predictive of final treatment response in 18 of the 22 patients with metastatic melanoma after two cycles of Ipilimumab (22). 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 to infer 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 FDG-PET might include the evaluation of other PET-based variables, our results suggest the potential capability of 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 ICI, it was demonstrated that as much as a half of patients showing residual disease on CT have negative FDG-PET scans (23-26). Indeed, the risk for PD misclassification after ICI was higher in initial observations, which found that 10% of patients that would have been misclassified as disease progression by World Health Organization criteria during ipilimumab for melanoma, showed clinical response (including PR and SD) (27). Subsequently, the rate of pseudoprogressions was reported from 0 to 6% in NSCLC treated with ICI and, at present, pseudoprogression should mainly be considered when the clinical condition of the patient in apparent radiological progression is concomitantly improving (28-29). In keeping with this evidence, recent FDG-PET studies introduced the evaluation of clinical benefit into the definition of response (25-26, 30). In this framework, our findings do not support the routine use of FDG-PET to monitor response to Nivolumab in all NSCLC patients, while they support the role of FDG-PET in the prognostic stratification of patients defined as PD according to CT-based criteria. Accordingly, FDG-PET evaluation might play a role in patients with suspected pseudoprogression although it should be underlined that only the availability of a baseline FDG-PET may allow to confirm or exclude progression by means of FDG-PET later in the course of the disease.

The present study has some drawbacks. First, the relatively small population. However, its monocentric nature represents one of its strengths. Accordingly, all the 48 enrolled patients were submitted to FDG-PET using the same PET/CT scanner avoiding the possible influence of inter-scanner variability on FDG-PET results. Second, the present preliminary findings apply to the PERCIST and imPERCIST-based assessment of FDG-PET and may not necessary be applied to other methods used to evaluate PET-based response to ICI. In fact, even in NSCLC patients treated with conventional chemotherapy, no consensus has been yet reached on the best PET-based approach to assess response to therapy and other methods based on other criteria such as EORTC criteria or metabolic tumor volume assessment (MTV) should also be considered. (30, 31). Indeed, Kaira et al addressed the role of several PET-derived parameters including metabolic tumor volume (MTV) and total lesion glycolysis (TLG) and highlighted the value of TLG 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 been to-date validated and in a recent position paper by Aide and colleagues the use of either the EORTC PET response criteria or PERCIST is

still suggested for the evaluation of FDG uptake changes in target lesion in patients treated with ICI (32). Moreover, in the recommendations by Aide and colleagues the addition to the evaluation of PET response of the computation of MTV and TLG is mentioned just as a possibility. Therefore, given the fact that the best method to compute MTV even in patients treated with conventional chemotherapy is still a matter of debate, in the present study we decided to privilege the exploration of several different CT and PET approaches based on more well-established measurements (i.e. the SULpeak). With respect to CT-based response evaluation, besides irRC criteria, the use of iRECIST criteria for CT-response evaluation in patients treated with ICI have more recently proposed. Actually, irRC were among the first criteria proposed and at the time of definition of design and approval of the present study, there were no other accepted or validated response criteria. The use of IrRC was thus part of the prospective design of the study thus obviously influencing recruited patients' management. The added value of PET-based criteria with respect to iRECIST was thus not addressed in the present study.

Another potential limitation of the present study is related to the multivariate analysis. In fact, we did not consider other potential prognostic factors in the cohort able to influence OS and consequently the impact of the different response criteria investigated. Finally, one of the binary classification 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, while 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 be thus considered with caution due to 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 FDG-PET in the general population of NSCLC patients treated with Nivolumab, it supports the prognostic potential of the metabolic response assessment and it confirms the importance of acquiring baseline FDG-PET data to compare with post-treatment examinations in complex cases potentially improving the therapeutic decision-making.

DISCLOSURES

Conflicts of interest: CG: Honoraria from Astra Zeneca, BMS, Boehringer Ingelheim, MSD, Roche; FG: Advisory Role: Ad Hoc Advisory Boards/Consultations (last 3 years), Eli Lilly, Roche, Boehringer Ingelheim, AstraZeneca, Pierre Fabre, BMS, MSD, Novartis Honoraria: Seminar/Talks to Industry (last 3 years), Eli Lilly, Roche, Boehringer Ingelheim, AstraZeneca, Pierre Fabre, AMGEN, Celgene, BMS, MSD, Research Funding (last 3 years); SM: Speaker honoraria from General Electric Healthcare (2017). No other potential conflicts of interest relevant to this article exist.

KEY POINTS

QUESTION: What is the role of FDG-PET in the evaluation of response to therapy in non-small-cell lung cancer (NSCLC) patients treated with immune checkpoint inhibitors?

PERTINENT FINDINGS: The present mono-institutional translational research trial strongly suggest the prognostic potential of the metabolic response assessment. Indeed, partial metabolic response assessed by PERCIST criteria predicted longer overall survival with respect to CT-based partial response.

IMPLICATIONS FOR PATIENT CARE: It might be useful to acquire baseline FDG-PET data in NSCLC patients' candidates to immune checkpoint inhibitors to compare with post-treatment examinations in complex cases potentially improving the therapeutic decision-making.

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;45:228-47.

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;19:3936-43.

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;15:7412-20.

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, Schöder H, Humm JL, Ni A, Michaud L, Nakajima R, Yamashita R, Wolchok JD, Weber WA. (18)F-FDG PET/CT for monitoring of ipilimumab therapy in patients with metastatic melanoma. J Nucl Med. 2019 Mar;60(3):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;373:123-35.

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-50S.

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;33:159-74.

18. Bergeri I, Michel R, Boutin JP. [Everything (or almost everything) about the kappa coefficient]. *Med Trop Rev Med Trop (Mars)*. 2002;62:634-6.

19. Lara PN Jr, Redman MW, Kelly K, Edelman MJ, Williamson SK, Crowley JJ, Gandara DR. 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;26:463-7

20. Borcoman E, Kanjanapan Y, Champiat S, Kato S, Servois V, Kurzrock R, Goel S, Bedard P, Le Tourneau C. Novel patterns of response under immunotherapy. *Ann Oncol.* 2019;30:385-396

21. Goldfarb L, Duchemann B, Chouahnia K, Zelek L, Soussan M. Monitoring anti-PD-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;7:27848-54.

23. Sachpekidis C, Anwar H, Winkler J, et al. The role of interim (18)F-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, Dimitrakopoulou-Strauss 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, Schöder 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;41:156-8.

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;15:7412-20.

29. Vrankar M, Unk M. Immune RECIST criteria and symptomatic pseudoprogression in non-small cell lung cancer patients treated with immunotherapy. *Radiol Oncol*. 2018;52:365–9.

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 (18)F-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;45:376–83.

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.

Figure titles and legends.

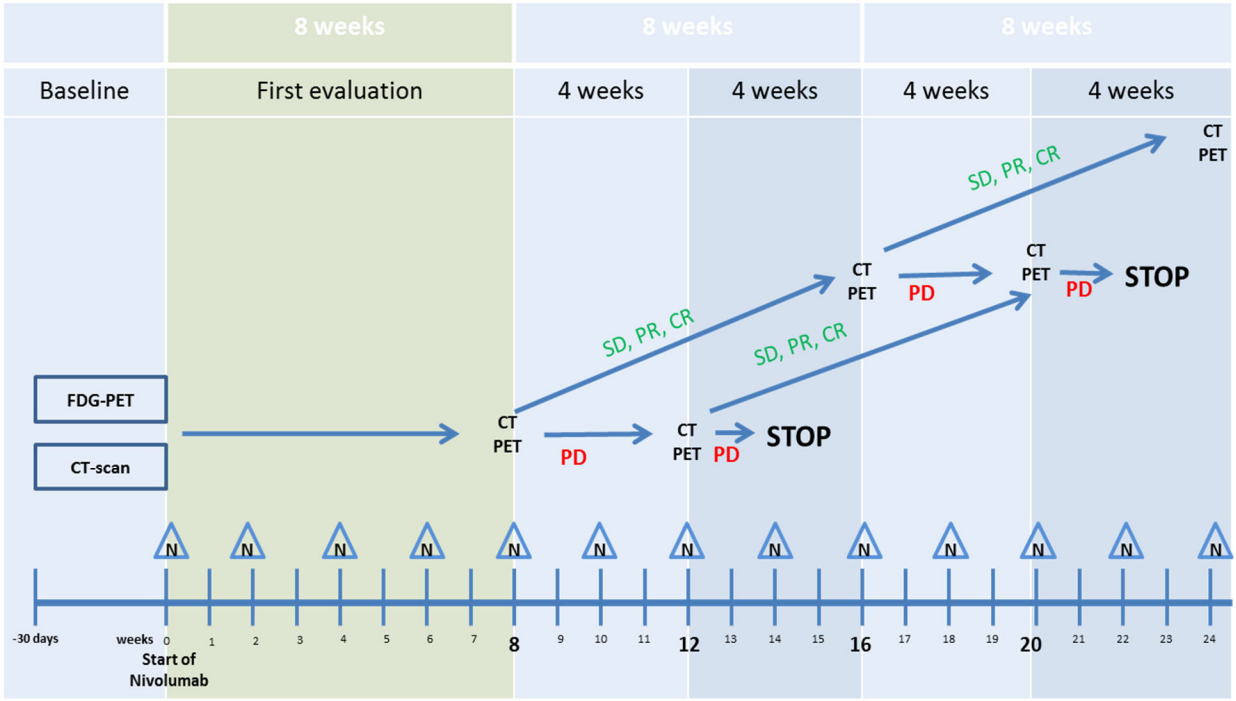

Figure 1. Schematic Representation of Study Design. first evaluation was performed with CT-scan and FDG-PET after 8 weeks (4 cycles of Nivolumab). If patient experienced a PD according to CT criteria, they repeated the evaluation after further 2 cycles (4 weeks). Otherwise, in case of SD, PR or CR evaluation of response was repeated every 4 cycles. CT: CT-Scan; PET: FDG-PET; PD: Progressive disease; SD: Stable disease; PR: Partial response; CR: Complete response.

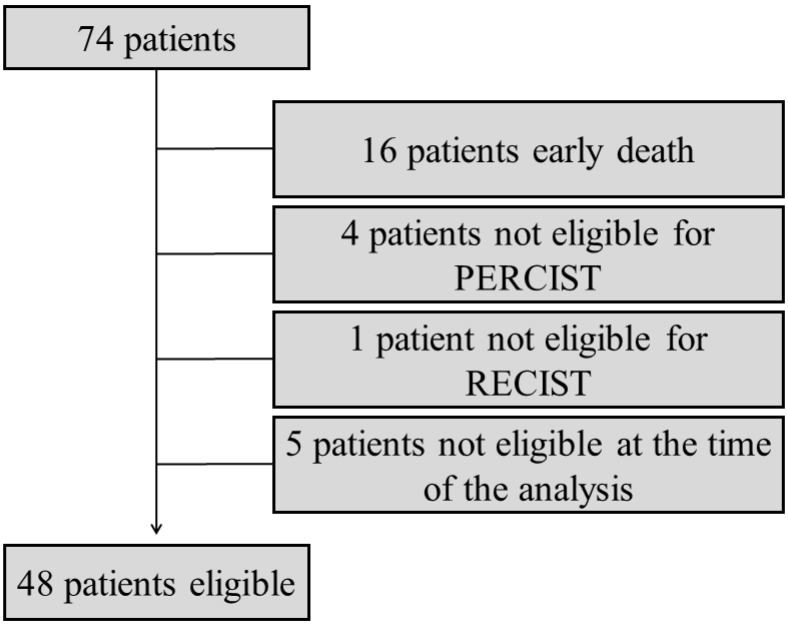

Figure 2: Patients eligible for the study.

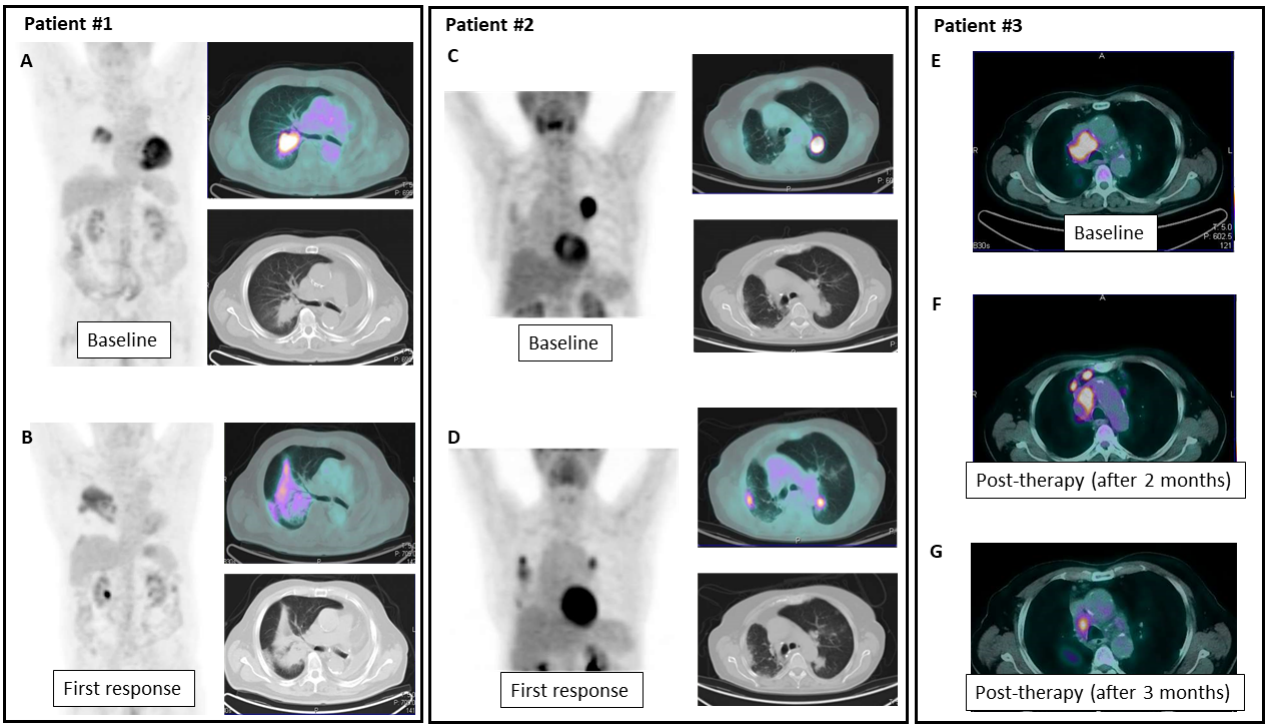

Figure 3: Representative PET/CT images of three patients showing discordant response at CT-based and PET-based evaluation or between PERCIST and imPERCIST criteria. Patient 1 was classified as PD with RECIST criteria due to the significant increase in the dimensions of the perihilar lesion in the right lung, however marked reduction of SULpeak allowed to classify this patient as PMR according to both PERCIST and imPERCIST criteria (panels A and B). On the contrary, Patient2 showed a marked lesions' shrinkage but (possibly due to inflammatory infiltration) lesions' metabolism was substantially stable (or even mildly increased in case of the right pleural metastasis). This patient was classified as SMD with both PERCIST and imPERCIST but response was more evident on CT images (Panel C and D). Patient3 was differently classified by PERCIST and imPERCIST criteria. In fact at first response (2 months after therapy) he was classified as SMD for imPERCIST criteria while due to the presence of new lesions he was considered as PERCIST-PMD. Of note after one further month new lesions disappeared and lesions present since baseline showed a significant reduction of both their dimensions and metabolism (panels E, F, G**).**

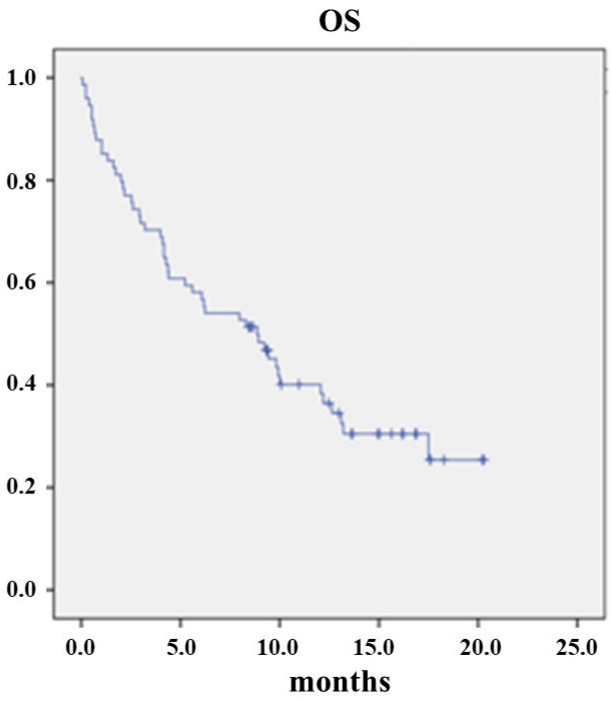

Figure 4: OS curves of the overall population of the study.

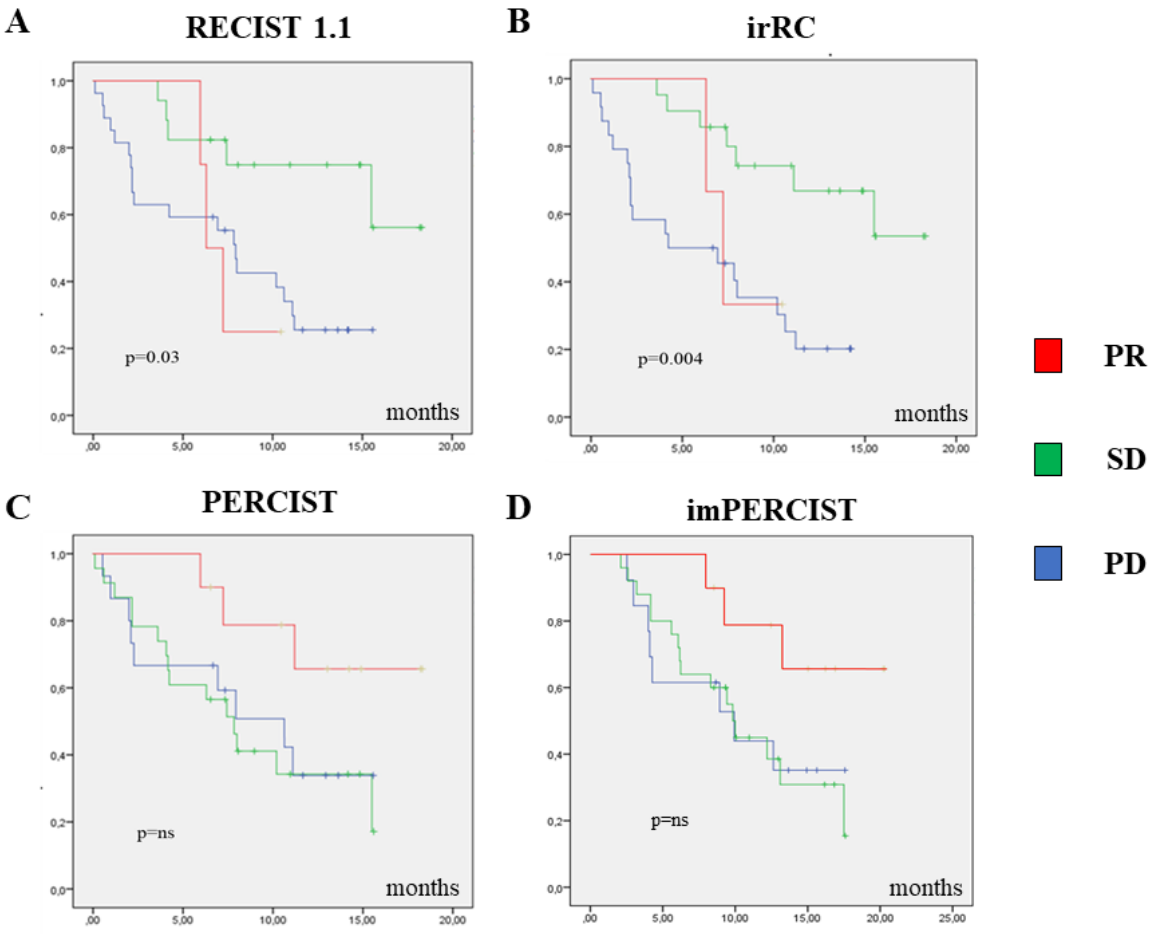

Figure 5: OS according to CT-based and PET-based criteria

Survival OS curves assessed by RECIST (a), irRC (b), PERCIST (c) and imPERCIST (d) -based response criteria at first evaluation. PD, Progressive Disease; SD, Stable Disease; PR, partial response;

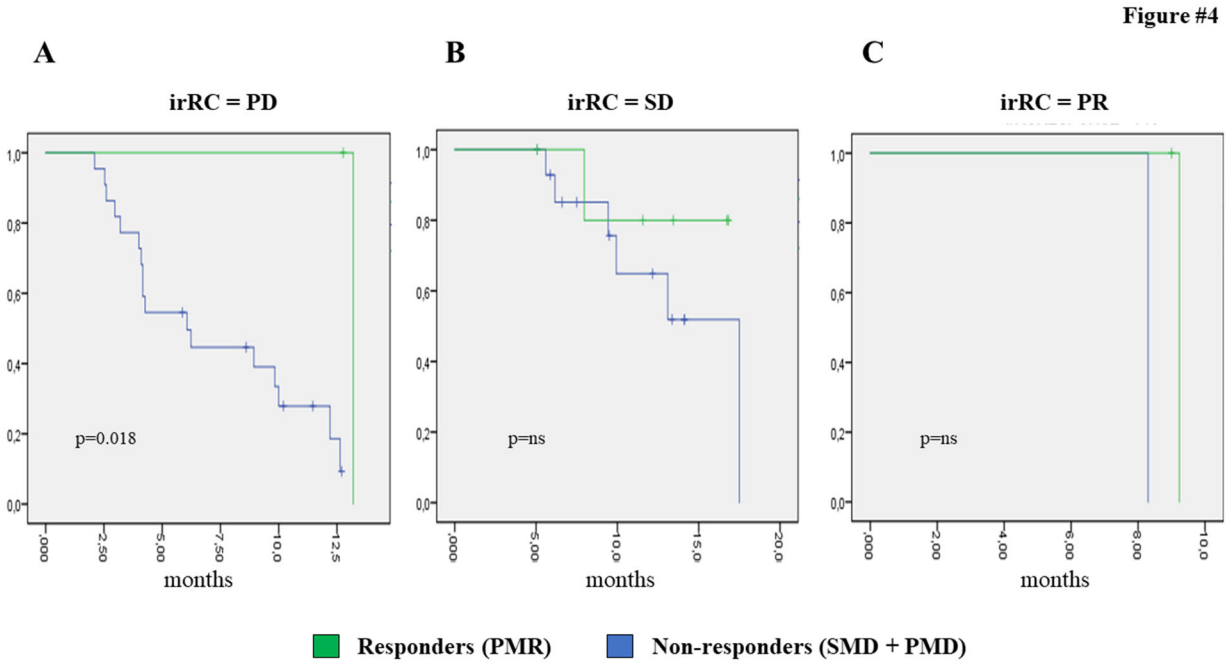

Figure 6: OS according to PET classes within irRC scores

Survival OS curves assessed by means of the binary classification of fist response according to both PET-based criteria in "responders" (PMR) or "non-responders" (SMD+PMD), in patients classified according to irRC classes (irRC-PD, irRC-SD and irRC-PR, respectively). PD, Progressive Disease; SD, Stable Disease; PR, partial response; PMD, Progressive Metabolic Disease; SMD, Stable Metabolic Disease; PMD, Progressive Metabolic Disease.

Table 1. Patient's characteristics.

Age		70 (range 44-85)
Gender	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%)
	Non-squamous	37/48 (77%)
Prior lines of therapy	1	17/48 (35%)
	2	13/48 (27%)
	≥ 3	18/48 (38%)

ECOG, Eastern Cooperative Oncology Group (ECOG) Performance Status

Table 2a. Concordance between RECIST 1.1 and PERCIST scores at first evaluation.

		PERCIST			
		PMD	SMD	PMR	Total
RECIST 1.1	PD	14	11	2	27
		29.2%	22.9%	4.2%	56.3%
	SD	1	11	5	17
		2.1%	22.9%	10.4%	35.4%
	PR	0	1	3	4
		0%	2.1%	6.3%	8.3%
Total		15	23	10	48
		31.3%	47.9%	20.8%	100%

PD, Progressive Disease; SD, Stable Disease; PR, partial response; PMD, Progressive Metabolic Disease; SMD, Stable Metabolic Disease; PMD, Progressive Metabolic Disease.

Table 2b. Concordance between irRC and PERCIST scores at first evaluation

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

PD, Progressive Disease; SD, Stable Disease; PR, partial response; PMD, Progressive Metabolic Disease; SMD, Stable Metabolic Disease; PMD, Progressive Metabolic Disease.

Table 3a. Concordance between RECIST 1.1 and imPERCIST scores at first evaluation.

		imPERCIST			
		PMD	SMD	PMR	Total
RECIST 1.1	PD	13	12	2	27
		27.1%	25%	4.2%	56.3%
	SD	0	12	5	17
		0%	25%	10.4%	35.4%
	PR	0	1	3	4
		0%	2.1%	6.3%	8.3%
Total		13	25	10	48
		27.1%	52.1%	20.8%	100%

PD, Progressive Disease; SD, Stable Disease; PR, partial response; PMD, Progressive Metabolic Disease; SMD, Stable Metabolic Disease; PMD, Progressive Metabolic Disease.

Table 3b. Concordance between irRC and imPERCIST scores at first evaluation

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

PD, Progressive Disease; SD, Stable Disease; PR, partial response; PMD, Progressive Metabolic Disease; SMD, Stable Metabolic Disease; PMD, Progressive Metabolic Disease.

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

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

1. Lung Cancer Unit, IRCCS Policlinico San Martino, Genoa, Italy

2. Nuclear Medicine Unit, IRCCS Policlinico San Martino, Genoa, Italy

3. Medical Oncology Unit, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy

4. Radiology Unit, IRCCS Policlinico San Martino, Genoa, Italy

5. Epidemiology Unit, IRCCS Policlinico San Martino, Genoa, Italy

6. Department of health sciences (DISSAL), University of Genoa, Italy

7. Nuclear Medicine Unit, Humanitas Clinical and Research Center - IRCCS, Rozzano, Italy

Corresponding author:

Prof. Silvia Morbelli, MD, PhD

Nuclear Medicine Unit

Department of health sciences (DISSAL), University of Genoa

IRCCS Policlinico San Martino, Genoa, Italy

Phone: +390105553038

e-mail: silvia.morbelli@unige.it

ORCID iD: https://orcid.org/0000-0002-7833-6095

Supplementary Materials.

Images acquisition protocol.

The CT scan parameters were set as follows. Arterial phase: slice thickness of 5 mm, pitch 0.8, tube rotation speed 0.5 s, 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. Portal phase: slice thickness 5 mm, pitch 0.8, tube rotation speed 0.5 s, 120 kV, reference 175 mA with the same modulation system; 2-mm thick slices 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 reviewing by the referring physician, while thinner slices were used for multiplanar imaging of vessels, bone (ribs and spine), and for high-resolution scanning of lung and liver lesions.

FDG-PET was performed according to the international guidelines (1) using a 16-slices PET/CT hybrid system (Biograph 16, Siemens Medical Solutions, Knoxville TN, USA). Briefly, patients fasted overnight prior to the intravenous administration of 300-400 MBq of FDG, which was performed in a quiet room, with the patient lying in a recumbent position and instructed not to move. Blood glucose was measured before tracer injection, as to ensure blood glucose levels <160 mg/dl. To minimize artifacts caused by the urinary tract, patients were asked to drink 500 mL of water 1h prior to image acquisition and to empty the bladder just before the acquisition start. Imaging started 60±15 minutes after intravenous tracer administration. The technical parameters of the 16-detector row, helical CT scanner included a gantry rotation speed of 0.5 s and 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 spacing of 3.75 mm. The trans-axial 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 emissive data and anatomical localization of PET dataset. Emissive scan was performed in 3D mode, shortly after CT acquisition, with a 3-min acquisition per bed position. PET sinograms were reconstructed by means of ordered-subset expectation maximization (OSEM) iterative reconstruction algorithm (three iterations, eight subsets). Scan was performed starting from the orbital plane on to the mid-thigh, except for the cases where the clinical history demanded a whole body, vertex-to-toes scan.

Definition of group response.

CT data were interpreted by physicians experienced in response evaluation with both RECIST 1.1 and irRC (GR, FG, GC), blinded to PET/CT results. Similarly, FDG-PET data were interpreted according to PERCIST criteria by physicians experienced in PERCIST-based response evaluation (MB, SM) blinded to CT results. Response evaluation criteria are detailed elsewhere (2-4) and briefly summarized below.

Definitions of response group according to RECIST 1.1: as: 1) complete response (CR: disappearance of all target lesions); 2) partial response (PR: at least 30% decrease in the sum of the longest diameter of five target lesions); 3) progressive disease (PD: at least a 20% increase in the sum of the longest diameter of five target lesions); and 4) stable disease (SD: neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD) compared to the baseline evaluation and previous one.

Definitions of response group according to IrRC: this evaluation requires that tumor burden was calculated and compared to the baseline and with the previous one. Tumor burden was defined as the sum of the products of the two largest perpendicular diameters of every index lesions (five lesions per organ, up to ten visceral lesions and five cutaneous index lesions). All new lesion do not define progressive disease, but have to be added to the tumor burden. According to these criteria: 1) complete response (CR: disappearance of all target lesions); 2) partial response (PR: at least a 50% decrease in tumor burden) 3) progressive disease (PD: at least a 25% increase in tumor burden); stable disease (SD neither decrease more than 50% of Tumor burden nor increase more than 25% of tumor burden).

Definition of group response according to PERCIST: FDG PET data were interpreted according to PERCIST criteria by physicians experienced in PERCIST-based response evaluation. PET readers were blinded to CT results. PERCIST recommends the use of lean body mass for SUV normalization (SUL). The background area was drawn as a 3-cm-diameter spheric ROI in the right lobe of the liver as defined in the criteria. The SULpeak of up to 5 lesions on the baseline and follow-up scan was summed (maximum of 2 per organ). Since the hottest lesions were selected in each scan, target lesions on follow-up scans were not necessarily the same as target lesions at baseline. The five lesions with highest SUL were identified, and a 1.2-cm-diameter spheric ROI was drawn in the hottest part of that lesions. The ROIs were placed in the area of the tumor where it resulted in the highest possible mean SUL (SULmean). SULmean of this ROI was defined as SULpeak. The baseline target lesions had to meet the PERCIST 1.0 definition of measurable lesions. The investigators checked that no other lesion could give a higher SULpeak. On subsequent scans, SULpeak could be located in a different lesion from the one measured at baseline, as long as the lesion had been present since baseline. In the follow-up PET/CT scans if SULpeak was decreasing, response was calculated as ΔSULpeak between baseline and actual follow-up divided

by baseline SULpeak X 100%. If SULpeak increased, response was calculated as ΔSULpeak between lowest registered and actual follow-up divided by lowest registered SULpeak x 100%. Response was classified on each scan according to the 4 categories defined in the criteria set. CMR was complete resolution of 18F-FDG uptake within all lesions to a level less than or equal to that of mean liver activity and indistinguishable from background blood-pool levels. PMR was a reduction of at least 30% in SULpeak and an absolute drop of 0.8 SULpeak units. PMD was an increase of at least 30% in SULpeak or a new 18F-FDG–avid lesion. SMD was between PMR and PMD.

Response according to imPERCIST was performed in the same way as described for PERCIST but following the modified method proposed by Ito and colleagues. In imPERCIST criteria the appearance of new lesions alone did not result in PMD and thus, PMD was defined only by an increase of the sum of SULpeaks by 30%. New lesions were included in the sum of SULpeak if they showed higher uptake than existing target lesions or if fewer than 5 target lesions were detected on the baseline scan.

References

1. E.A. Eisenhauer; P. Therasse; J. Bogaerts; L.H. Schwartz; D. Sargent; R. Ford; J. Dancey; S. Arbuck; S. Gwyther; M. Mooney; L. Rubinstein; L. Shankar; L. Dodd; R. Kaplan; D. Lacombe; J. Verweij (2009), "New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1)", European Journal of Cancer, 45 (2): 228–247
2. 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;15:7412–20
3. 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 May;50 Suppl 1:122S-50S.
4. O JH, Lodge MA, Wahl RL. Practical PERCIST: A Simplified Guide to PET Response Criteria in Solid Tumors 1.0. Radiology. 2016 Aug;280(2):576-84.

5. Ito K, Teng R, Schöder H, Humm JL, Ni A, Michaud L, Nakajima R, Yamashita R, Wolchok JD, Weber WA. (18)F-FDG PET/CT for Monitoring of Ipilimumab Therapy in Patients with Metastatic Melanoma. J Nucl Med. 2019 Mar;60(3):335-341.