Received: 17 January 2023 | Revised: 10 April 2023 | Accepted: 28 April 2023

DOI: 10.1111/jne.13311

ORIGINAL ARTICLE

Journal of Neuroendocrinology WILEY

Proposal of early CT morphological criteria for response of liver metastases to systemic treatments in gastroenteropancreatic neuroendocrine tumors: Alternatives to RECIST

Louis de Mestier[1] | Matthieu Resche-Rigon[2] | Clarisse Dromain[3] | Angela Lamarca[4] | Anna La Salvia[5] | Lesley de Baker[6] | Uli Fehrenbach[7] | Sara Pusceddu[8] | Annamaria Colao[9,10] | Ivan Borbath[11] | Robbert de Haas[12] | Maria Rinzivillo[13] | Alessandro Zerbi[14] | Luigi Funicelli[15] | Wouter W. de Herder[16] | Andreas Selberherr[17,18] | Anna Dorothea Wagner[19] | Prakash Manoharan[20] | Andrea De Cima[21] | Willem Lybaert[22] | Henning Jann[23] | Natalie Prinzi[8] | Antongiulio Faggiano[9] | Laurence Annet[24] | Annemiek Walenkamp[25] | Francesco Panzuto[13,26] | Vittorio Pedicini[27] | Maria Giovanna Pitoni[28] | Alexander Siebenhuener[29] | Marius E. Mayerhoefer[30,31] | Philippe Ruszniewski[1] | Marie-Pierre Vullierme[32]

Correspondence
Marie-Pierre Vullierme, Department of Radiology, Beaujon Hospital, 100 boulevard du Général Leclerc, 92110 Clichy, France.
Email: mpvullierme@gmail.com

Abstract

RECIST 1.1 criteria are commonly used with computed tomography (CT) to evaluate the efficacy of systemic treatments in patients with neuroendocrine tumors (NETs) and liver metastases (LMs), but their relevance is questioned in this setting. We aimed to explore alternative criteria using different numbers of measured LMs and thresholds of size and density variation. We retrospectively studied patients with advanced pancreatic or small intestine NETs with LMs, treated with systemic treatment in the first-and/or second-line, without early progression, in 14 European expert centers. We compared time to treatment failure (TTF) between responders and non-responders according to various criteria defined by 0%, 10%, 20% or 30% decrease in the sum of LM size, and/or by 10%, 15% or 20% decrease in LM density, measured on two, three or five LMs, on baseline (≤1 month before treatment initiation) and first revaluation (≤6 months) contrast-enhanced CT scans. Multivariable Cox proportional hazard models were performed to adjust the association between response criteria and TTF on prognostic factors. We included 129 systemic

For affiliations refer to page 10

treatments (long-acting somatostatin analogs 41.9%, chemotherapy 26.4%, targeted therapies 31.8%), administered as first-line (53.5%) or second-line therapies (46.5%) in 91 patients. A decrease ≥10% in the size of three LMs was the response criterion that best predicted prolonged TTF, with significance at multivariable analysis (HR 1.90; 95% CI: 1.06–3.40; $p = .03$). Conversely, response defined by RECIST 1.1 did not predict prolonged TTF ($p = .91$), and neither did criteria based on changes in LM density. A ≥10% decrease in size of three LMs could be a more clinically relevant criterion than the current 30% threshold utilized by RECIST 1.1 for the evaluation of treatment efficacy in patients with advanced NETs. Its implementation in clinical trials is mandatory for prospective validation. Criteria based on changes in LM density were not predictive of treatment efficacy.

Clinical Trial Registration: Registered at CNIL-CERB, Assistance publique hopitaux de Paris as "E-NETNET-L-E-CT" July 2018. No number was assigned. Approved by the Medical Ethics Review Board of University Medical Center Groningen.

KEYWORDS
computed tomography, neuroendocrine tumors, response evaluation, systemic treatments

1 | INTRODUCTION

Digestive well-differentiated neuroendocrine tumors are relatively rare neoplasms with increasing incidence.[1] While they can arise from all the digestive system, gastroenteropancreatic neuroendocrine tumors (GEP-NETs) are among the most frequent localizations. Most small intestine and pancreatic NETs are associated with distant metastases, mainly liver metastases (LMs).[2]

Because surgical resection of LMs is rarely feasible with curative intent, almost all patients with NET-associated LMs receive systemic therapies.[2–4] During the last decades, several therapies have demonstrated antitumor efficacy in randomized controlled trials performed in this setting, including long-acting somatostatin analogs (SSAs) (lanreotide and octreotide),[5,6] everolimus,[7,8] sunitinib in pancreatic NETs[9] and peptide-radionuclide receptor therapy (PRRT) in midgut NETs.[10] In addition, chemotherapy, predominantly based on alkylating-based combinations, remains a cornerstone of treating NET-associated LMs, mainly from pancreatic origin.[11,12] Evaluation of the efficacy of systemic treatments for GEP-NETs currently relies on Response Evaluation Criteria in Solid Tumors (RECIST 1.1),[13] both for daily practice and clinical studies.[4,14] The RECIST-defined objective response consists in a reduction above or equal to 30% of the sum of the greater diameter of two target lesions as documented on computed tomography (CT) scans performed during follow-up, in comparison with baseline scans.[13]

However, RECIST criteria have not been specifically validated in patients with metastatic GEP-NETs and are challenged in this setting.[14–16] First, the optimal percent of decrease in the size of NET-associated LMs to define response has not been properly explored. GEP-NETs are characterized by generally slow growth, highly vascularized lesions and the fact that most treatments may have more cytostatic than cytotoxic antitumor effects. Hence, the 30% threshold of size variation defining objective response according to RECIST criteria is rarely reached, contrasting with generally prolonged tumor control.[16,17] In addition, the optimal number of target LMs to be measured has not been clearly defined. Because different LMs can often have discordant evolution, the relevance of considering only two target lesions can be questioned. Finally, most NET-associated LMs typically show high vascularization. Treatments (mainly targeted therapy) frequently induce a decrease in tumor density measured on CT, which is thought to be related to LM necrosis. Density-related response criteria, initially reported for gastrointestinal stromal tumors (Choi criteria) and hepatocellular carcinoma (mRECIST),[18,19] might have a role in the evaluation of NET response to systemic therapies.[15]

We aimed to explore criteria alternative to RECIST 1.1, using different numbers of measured LMs and thresholds of size and density variation, for the early evaluation of objective response to systemic therapies in patients with advanced pancreatic or small intestine NETs and LMs.

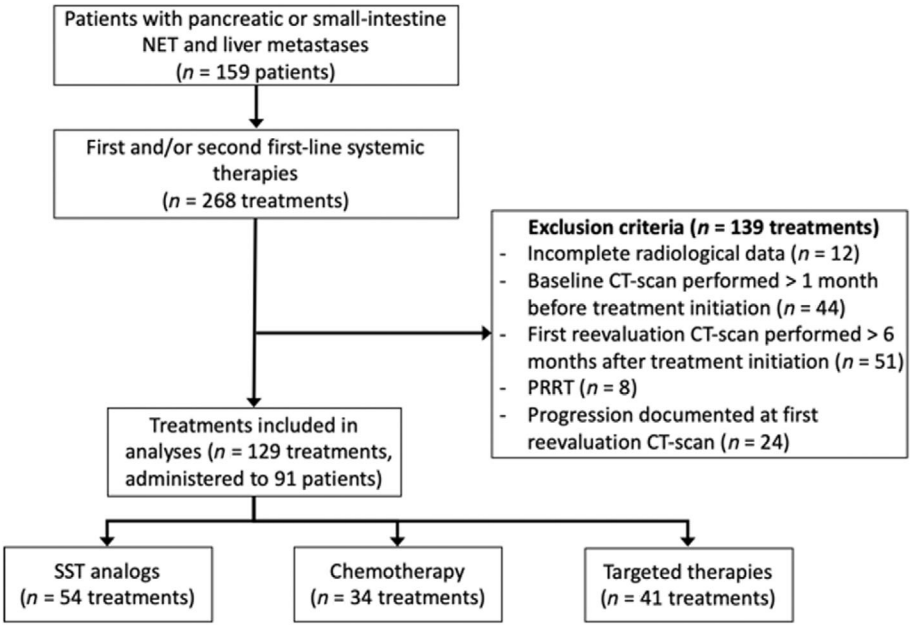
2 | METHODS

2.1 | Patients

We conducted a retrospective, international study under the auspices of the European Neuroendocrine Tumor Society (ENETS), in 14 centers specialized in the management of NETs among 50 centers initially solicited (Figure 1).

Each participating center was asked to include 12 patients. Patients could be included if they had a sporadic NET of pancreatic or small intestine origin, associated with LMs measurable on CT, and treated with at least one first-line systemic therapy. All GEP-NETs had to be histologically proven, well-differentiated and of any grade (G1, G2

FIGURE 1 Flow chart of the study.



or G3) by local review, following pathological consensus guidelines.[20,21] Patients with neuroendocrine carcinoma and NET from other primary sites and those with genetic predisposition were not eligible. The choice of patients to include was left at the discretion of local investigators, with no obligation to include consecutive patients.

The systemic treatments eligible for analyses included SSAs (lanreotide and octreotide), everolimus, sunitinib, chemotherapy or PRRT, received in the first- or second-line. Treatment dose reduction up to 30% (or 50% for everolimus) was allowed. Patients who received any locoregional treatment (such as transarterial embolization or transcutaneous ablation) before or during systemic therapy were excluded. Of note, observations with treatments using PRRT were secondarily excluded due to the limited number of patients ($n = 8$).

A minimal number of two LMs monitored was mandatory for analyses. The monitoring of treatment efficacy had to rely on contrast-enhanced CT with arterial and portal phases, performed at baseline and during regular follow-up. Observations with systemic treatments were excluded from the analyses in case radiological data were either incomplete or did not rely on CT, if baseline (pretherapeutic) CT was performed more than one month before treatment initiation, or if the first evaluation CT was performed more than six months after baseline imaging or showed early tumor progression by RECIST 1.1.

This study was performed according to the Helsinki convention. The study protocol of this retrospective medical research was approved by the Medical Ethics Review Board of University Medical Center Groningen, and ethics or audit committees at each institution, wherever required.

2.2 | Data collection

Anonymized data were retrospectively collected in each center, including primary tumor origin, ECOG performance status, functioning syndrome, Ki-67 index and tumor grade according to the 2019 WHO classification.[20] Percent of metastatic liver involvement was classified as 1%–10%, 11%–25%, 26%–50% or >50%.[22] The beginning and end dates of each treatment, and the reason for treatment discontinuation (progression, toxicity) were collected. We collected the date of treatment failure, defined in local expert tumor boards of each center, according to clinical, biological and morphological criteria per RECIST 1.1.

2.3 | CT analysis

All CT scan images were reviewed locally by experienced radiologists, using the same standardized analysis protocol. It was decided not to centralize morphological evaluation in order to reflect real-life setting, and because all participating institutions were high-volume ENETS Centers of Excellence involving highly specialized radiologists.

At least two, and ideally up to five target LMs, were defined on baseline CT scans for each patient. Target lesions were selected based on their size (those with the longest diameter; the minimal size of 10 mm recommended) and their suitability for accurate repeated measurements (the more clearly delineated and easy to locate in the liver) and were numbered by order of increasing suitability (i.e., T1 to T5). The largest diameter (mm) of each target lesion, as well as the density (in Hounsfield units [HU]) of the two main target lesions (T1 and T2) were measured on each CT performed at baseline and during the follow-up under the first- and second-line systemic therapy. LM density was measured on both arterial and portal venous phases, on the same CT slice as the one used for size measurement, using a manually drawn region of interest covering the whole lesion (i.e., not restricted to any central, hypodense portion).

2.4 | Statistical analyses

We defined various objective response criteria, considering different numbers of target LMs (2, 3 or 5), thresholds of percent decrease in the sum of the size of the target LMs (0%, 10%, 20% or 30%), and

TABLE 1 Characteristics of the 129 treatment lines included in the analyses.

Treatment line, *n* (%)	
First-line	69 (53.4)
Second-line	60 (46.5)
Type of treatment, *n* (%)	
Long-acting somatostatin analogs	54 (41.9)
Chemotherapy	34 (26.4)
Everolimus	32 (24.8)
Sunitinib	9 (6.98)
Number of cycles, median (IQR)	10 (6–14.8)
Dose reduction, *n* (%)	18 (13.9)
Primary NET, *n* (%)	
Small intestine	57 (44.1)
Pancreas	72 (55.8)
Functioning NET, *n* (%)	42 (32.5)
Carcinoid syndrome	32 (24.8)
Insulinoma	2 (1.6)
ViPoma	2 (1.6)
Other	6 (4.7)
Performance status at baseline	
PS-0	77 (59.7)
PS-1	52 (40.3)
Ki-67 index (%), median (IQR)[a]	5 (2–10)
WHO grade 2019, *n* (%)	
Grade 1 (Ki-67 < 3%)	40 (31.0)
Grade 2 (3 ≤ Ki-67 ≤ 20%)	77 (59.7)
Grade 3 (Ki-67 > 20%)	8 (6.2)
Not available	4 (3.1)
Liver metastatic involvement, *n* (%)	
0%–10%	42 (32.6)
11%–25%	28 (21.7)
26%–50%	21 (16.3)
>50%	21 (16.3)
Not available	17 (13.2)

[a]Four missing values.

percent decrease in LM density (10%, 15% or 20%, measured as the mean of the density of 2 LMs on arterial or portal phase of contrast-enhanced CT). Those response criteria were measured on the CT corresponding to the first evaluation (within 6 months following treatment initiation), compared to the baseline CT, for each distinct systemic treatment line.

Our primary endpoint was time to treatment failure (TTF), defined as the time elapsed between the first evaluation and the failure of systemic treatment. Treatment failure was assessed by the local expert tumor boards of each center according to the same guidelines using clinical and biological criteria, and morphological evolution of hepatic but also extra-hepatic disease (which was not measured specifically for this study, hence hindering assessment of progression-free survival) per RECIST 1.1.

Continuous and categorical variables were described by calculating their medians (interquartile 25–75) and frequencies (percentages), respectively. TTF according to each criterion was estimated using the Kaplan–Meier method and was described by median with 95% confidence interval. TTF was compared between patients classified as "responders" or "non-responders" according to each objective response criterion using log-rank tests and univariable Cox proportional hazard models.

The impact on TTF of OR criteria with a *p*-value <.10 in univariable analyses was further explored in multivariable Cox models, adjusted for primary NET (pancreas vs. small bowel), type of systemic treatment (SSAs, chemotherapies or targeted therapies), treatment line (first- or second-line), presence of functioning syndrome and tumor grade.

In addition, univariable sensitivity analyses were performed for treatments with five LMs measured only. Moreover, subgroup analyses corresponding to the three types of systemic treatment (SSAs, chemotherapies or targeted therapies) were performed.

3 | RESULTS

3.1 | Patients and treatments included in the study

Among 50 expert centers initially solicited, 23 did not answer and 13 others declined because patient follow-up was routinely performed using magnetic resonance imaging (MRI), or because previously performed imaging data was not available for local review. Among the 14 participating centers, 159 patients with pancreatic or small intestine NET and LMs were screened for eligibility, accounting for 268 different treatment lines. Among them, 139 treatment lines were excluded, mainly because baseline or first evaluation CT were performed too early or too late regarding treatment initiation, respectively (95/268, 35.4%) (Figure 1).

Overall, 129 systemic treatments, performed at first-line (53.5%) or in the second-line (46.5%), were included in the analyses. These treatments consisted of long-acting somatostatin analogs, chemotherapy or targeted therapies in 54 (41.9%), 34 (26.4%) and 41 (31.8%) of cases, respectively (Table 1). The median number of cycles administered was 10, with dose reductions in 13.9% of cases. Primary NET originated from the pancreas or small intestine in 55.8% and 44.1% of cases, respectively, and was associated with a functioning syndrome in 32.5% of cases. Median Ki-67 was 5% and most GEP-NETs were classified as G2 (59.7%) or G1 (31%). Metastatic liver involvement was above 25% in 32.6% of cases.

3.2 | Treatment discontinuation and TTF

The most frequent reasons for treatment discontinuation were progression ($n = 96$, 74.4%), therapeutic pause ($n = 13$, 10.1%) and unacceptable toxicity ($n = 10$, 7.8%). Overall, treatment failure occurred in

83.7% of cases. Median TTF from the first evaluation CT was 9.5 months (95% CI: 7.2–11.9) (Figure 2). The 6-, 12- and 24-month rates of survival without treatment failure from the first evaluation were 66% (95% CI: 58.2–74.8), 38.1% (95% CI: 30.2–48.2) and 17.8% (95% CI: 11.7–27.0), respectively.

The median global TTF from treatment initiation was 12.5 months (95% CI: 10.5–15.1). The 6-, 12- and 24-month rates of survival without treatment failure from treatment initiation were 82.8% (95% CI: 76.5–89.6), 52.4% (95% CI: 44.3–62.1) and 24.4% (95% CI: 17.5–34.1), respectively.

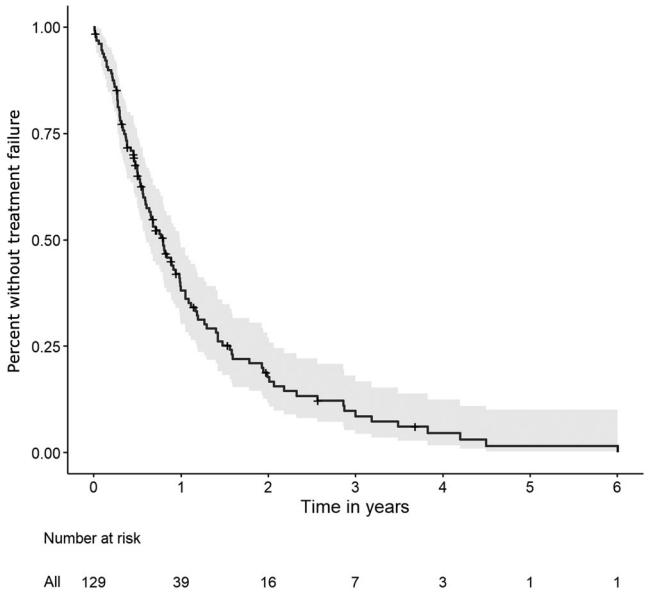


FIGURE 2 Estimation of time to treatment failure achieved by 129 first- or second-line systemic treatments administered in 91 patients with pancreas or small intestine neuroendocrine tumors and liver metastases.

3.3 | Association of size-based OR criteria with TTF

Objective response was assessed by measuring the variation in size of two, three or five LMs, for 129 (100%), 103 (79.8%) and 92 (71.3%) distinct systemic treatments, respectively (Table 2). The median size of target lesions was 23 mm (IQR, 16–38), with 96% of lesions measuring 1 cm or above.

When measuring two LMs (T1 and T2), responders defined by a decrease in size ≥30% between baseline and first evaluation CT (corresponding to RECIST 1.1 definition) were not associated with longer TTF compared to non-responders ($p = .91$) (Figure 3A). Moreover, responders defined by a decrease in size ≥20% or ≥10% of two LMs between baseline and first evaluation CT were not associated with more prolonged TTF compared to non-responders ($p = .41$ and $p = .35$, respectively), while responders defined by a decrease in size ≥0% were associated with a nonsignificant improvement of TTF ($p = .09$) (Figure 3B).

When measuring three LMs (T1–T3), responders defined by a decrease in size ≥30% or ≥20% were not associated with more prolonged TTF compared to non-responders ($p = .43$ and $p = .88$, respectively), while responders defined by a decrease in size ≥10% or ≥0% were associated with a nonsignificant improvement of TTF ($p = .10$ and $p = .052$, respectively) (Figure 3C,D).

Similarly, when measuring five LMs (T1–T5), responders defined by a decrease in size ≥30% or 20% were not associated with more prolonged TTF compared to non-responders ($p = .34$ and $p = .70$, respectively), while responders defined by a decrease in size ≥10% or ≥0% were associated with a nonsignificant improvement of TTF ($p = .13$ and $p = .0502$, respectively) (Figure 3E,F).

As more than two LMs could not be measured in all patients, these analyses were also performed in the subgroup of treatments with five LMs measured, as a sensitivity analysis (Table S1). Similar results were found, that is, RECIST criteria were not associated with TTF, while the strongest

TABLE 2 Association (hazard ratio) of various criteria of objective response based on the variation of the size of two, three or five liver metastases, with time to treatment failure for 129 systemic treatments.

Definition of objective response	Number of liver metastases measured		
	$n = 2$ ($n = 129$)	$n = 3$ ($n = 103$)	$n = 5$ ($n = 92$)
Decrease in size ≥ 30%(RECIST[13])	HR 0.96 95% CI: 0.42–2.19 $p = .91$	HR 1.50 95% CI: 0.55–4.11 $p = .43$	HR 1.76 95% CI: 0.55–5.59 $p = .34$
Decrease in size ≥ 20%	HR 1.35 95% CI: 0.66–2.79 $p = .41$	HR 0.95 95% CI: 0.49–1.84 $p = .88$	HR 1.16 95% CI: 0.56–2.41 $p = .70$
Decrease in size ≥ 10%	HR 1.25 95% CI: 0.79–1.97 $p = .35$	HR 1.59 95% CI: 0.92–2.75 $p = .10$	HR 1.61 95% CI: 0.87–2.99 $p = .13$
Decrease in size ≥0%	HR 1.40 95% CI: 0.95–2.07 $p = .09$	HR 1.54 95% CI: 1.00–2.39 $p = .052$	HR 1.60 95% CI: 1.00–2.57 $p = .0502$

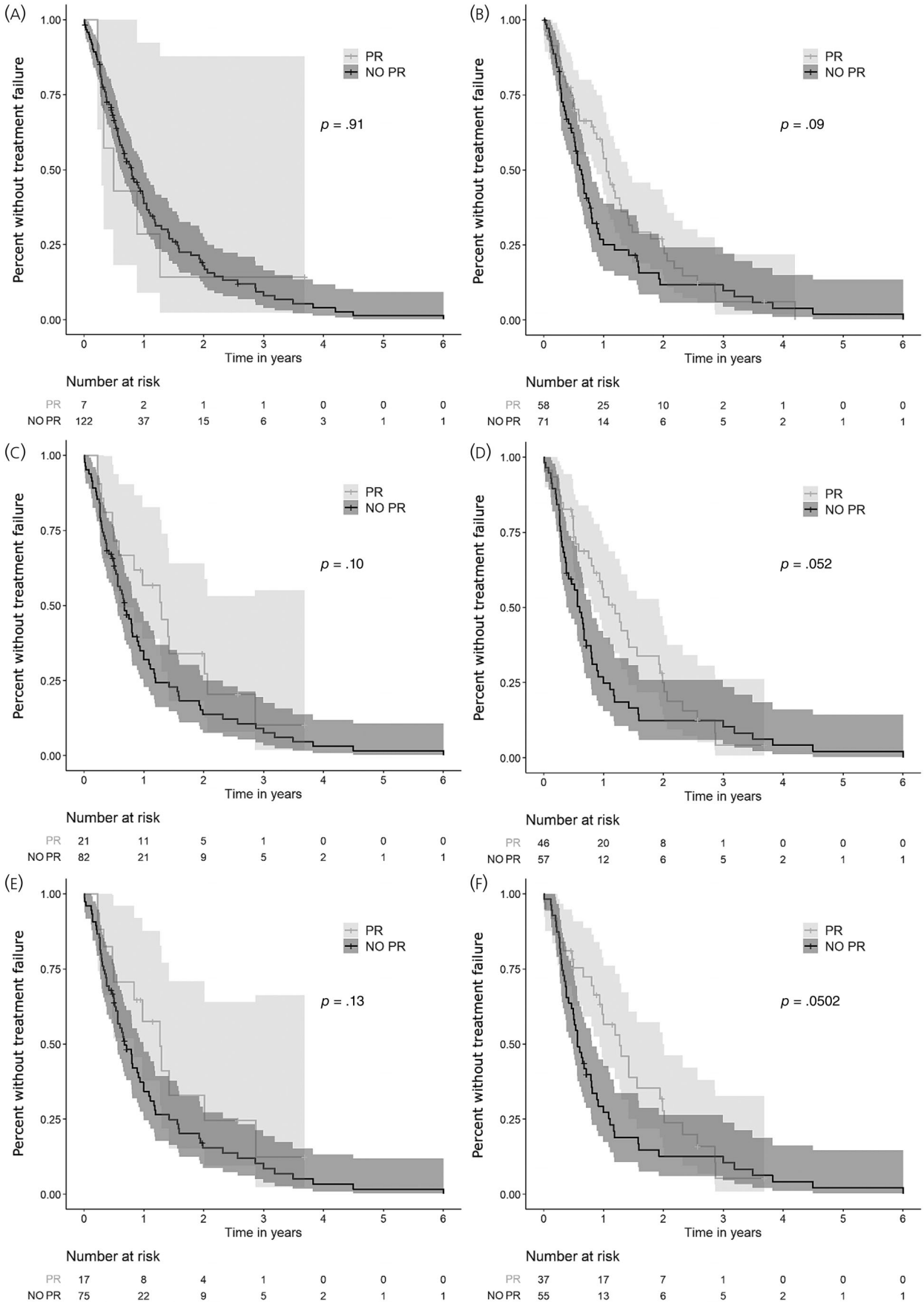


FIGURE 3 Legend on next page.

TABLE 3 Multivariable Cox proportional hazard analysis exploring the association between criteria of response and the risk of treatment failure.

	Multivariate analysis		
Definition of objective response	**HR (for absence of response)**	**95% CI**	***p*-value**
Decrease in size ≥0%, measured on 2 LMs	1.49	0.98–2.25	.06
Decrease in size ≥ 10%, measured on 2 LMs	1.34	0.82–2.18	.24
Decrease in size ≥0%, measured on 3 LMs	**1.66**	**1.03–2.70**	**.04**
Decrease in size ≥ 10%, measured on 3 LMs	**1.90**	**1.06–3.40**	**.03**
Decrease in size ≥0%, measured on 5 LMs	1.65	0.98–2.81	.06
Decrease in size ≥ 10%, measured on 5 LMs	**2.10**	**1.04–4.21**	**.04**

Note: Each response criteria were tested in one separate multivariable Cox model, adjusted for primary NET (pancreas vs. small intestine), type of systemic treatment (chemotherapy vs. targeted therapy vs. somatostatin analogs), treatment line (first- vs. second-line), presence of functioning syndrome and tumor grade (G1 vs. G2 vs. G3). Bold values denote statistical significance at the $p < 0.05$ level.
Abbreviations: HR, hazard ratio; LM, liver metastases.

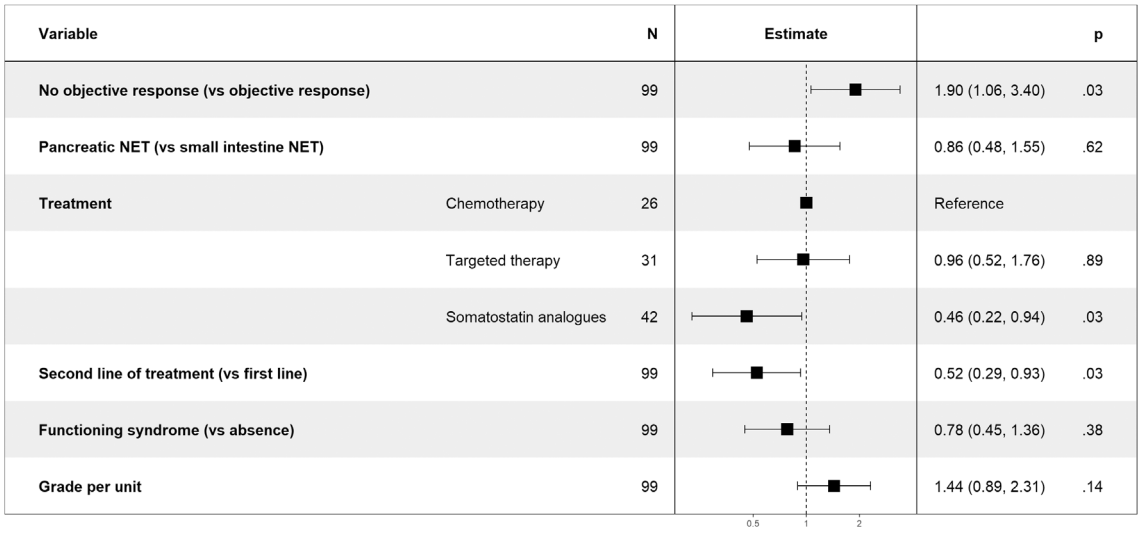


FIGURE 4 Forest plot of the multivariate analysis of time to treatment failure. After adjustment on potentially cofounding factors, the alternative response criteria (decrease in size ≥10%, measured on 3 liver metastases) was independently associated with significantly prolonged time to treatment failure. PR, partial response.

associations with TTF remained for decrease in size ≥10% or ≥0%, measured on three ($p = .09$ and $p = .04$, respectively) liver metastases.

3.4 | Multivariable analyses

On multivariable Cox proportional hazard analyses, adjusted for primary NET, type of systemic treatment, treatment line, presence of functioning syndrome and tumor grade, three objective response criteria were significantly associated with the risk of treatment failure: decrease ≥10% in the size of three LMs (HR 1.90; 95% CI: 1.06–3.40; $p = .03$), decrease ≥10% in the size of five LMs (HR 2.10; 95% CI: 1.04–4.21; $p = .04$) and decrease ≥0% in the size of three LMs (HR 1.66; 95% CI: 1.03–2.70, $p = .04$) (Table 3). The criteria defined by decrease ≥10% in the size of three LMs had the strongest statistical significance (Figure 4).

3.5 | Association of density-based response criteria with TTF

Because NETs are classically highly vascularized neoplasms, we evaluated the relevance of density-based response criteria, by measuring

FIGURE 3 Estimation of time to treatment failure in responders and non-responders according to various objective response criteria, defined as a variation in tumor size in comparison with baseline. (A) ≥30% measured on two liver metastases (RECIST 1.1 criteria); (B) ≥0% measured on two liver metastases; (C) ≥10% measured on three liver metastases; (D) ≥0% measured on three liver metastases; (E) ≥10% measured on five liver metastases; (F) ≥0% measured on five liver metastases.

TABLE 4 Association (hazard ratio) of various criteria of objective response based on the variation of density and size of two liver metastases with time to treatment failure for 118 systemic treatments.

Definition of objective response	Phase of contrast enhancement	
	Arterial ($n = 115$)	Portal ($n = 118$)
Decrease in density ≥ 20%	HR 0.94 95% CI: 0.56–1.57 $p = .80$	HR 1.01 95% CI: 0.44–2.30 $p = .99$
Decrease in density ≥ 10%	HR 0.86 95% CI: 0.56–1.31 $p = .47$	HR 1.07 95% CI: 0.68–1.69 $p = .92$
Decrease in density ≥ 15% and/or decrease in size ≥ 10% (Choi criteria[18])	HR 0.90 95% CI: 0.59–1.37 $p = .92$	HR 0.97 95% CI: 0.65–1.45 $p = .88$

the variation in density of two LMs (T1 and T2), on arterial phase or portal venous phase, for 115 (89.1%) and 118 (94.4%) distinct systemic treatments, respectively (Table 4).

Responders defined by a decrease in tumor density ≥20% or ≥10% were not associated with more prolonged TTF compared to non-responders, whenever it was measured on either arterial phase ($p = .80$ and $p = .47$, respectively) or portal phase ($p = .99$ and $p = .92$, respectively).

Similarly, responders defined by a decrease in LM density ≥ 15% and/or a decrease in size ≥10% were not associated with longer TTF compared to non-responders, whenever it was measured on either arterial phase ($p = .90$) or portal phase ($p = .60$, corresponding to Choi criteria[18]).

4 | DISCUSSION

In this study, we highlight that RECIST 1.1 criteria may not be appropriate to measure the response of NET-associated LMs treated with systemic therapies. Otherwise, we propose alternative criteria for objective response, defined as a decrease ≥10% in the size of three LMs. This cutoff did impact TTF, even when adjusted to other prognostic factors. Conversely, response criteria based on density variation were not predictive of treatment efficacy.

The original RECIST 1.0 criteria stated that evaluation of tumor burden and evolution would be measured by the minimal unidimensional size of up to 10 measurable lesions, with a maximum of five per organ site.[23] In 2009, revised RECIST 1.1 criteria proposed reducing the maximum number of lesions to measure from five to two per organ.[13] These criteria were proposed, and are used in clinical practice and research, for all types of neoplasms but do not take NET-related specific features into account, notably their slow evolution.[15,24] Due to the generally prolonged survival of patients, and the low rates of resectability of all diseases, most systemic treatments are administered with an intent of prolonging survival rather than achieving tumor debulking.[16] Accordingly, patient outcomes may not correlate with tumor shrinkage. Hence, the size variation of target lesions may not necessarily reach 30% in order to define response to treatment. This was underlined in one recent monocentric retrospective study, which reported that RECIST-defined response criteria could not predict time to progression in patients with advanced G1–G2 GEP-NETs treated with either chemotherapy, PRRT or everolimus.[17] There was even a trend toward shorter time to progression for those who achieved an objective response, which is concordant with our results (Figure 3A).

In our study, a decrease of 10% in LM size was the most robust threshold to distinguish responders from non-responders, measured on either three or five LMs. Consistently, this 10% of size variation threshold was previously reported to be the most optimal response criterion to predict progression-free survival in patients with advanced pancreatic NETs treated with sunitinib.[25] In this post hoc analysis of 237 patients included in the phase II and phase III studies of sunitinib, this threshold yielded the highest rate of correctly classified patients (67%). It had significant impact on progression-free survival ($p = .04$), while RECIST criteria did not ($p = .20$), which was confirmed in the multivariable analysis. Overall, a reduction of 10% in tumor size seems to be an accurate surrogate for treatment efficacy in GEP-NETs. As an alternative to evaluating variations in tumor size, assessment of tumor growth rate has shown very promising results. It may be especially suited for NETs as they are generally slow-growing. Indeed, the pre-therapeutic tumor growth rate may predict treatment efficacy, and its variation under treatment has been reported as a very interesting alternative to RECIST in the setting of NETs.[26]

We also aimed to define the most appropriate number of LMs to be measured for response evaluation. The RECIST 1.1 criteria state that a maximum of two lesions per organ and five in total should be measured.[13] However, these criteria may be difficult to apply in patients with NET-associated LMs, due to changes in their appearance following contrast administration, as well as the coalescence of lesions and the subsequent inability to delineate individual masses. In addition, paradoxical evolution of LMs, with some increasing and some decreasing in size on the same scan, is not rare. In our study, the measurement of three LMs was the most robust number to distinguish responders from non-responders, with thresholds of size variation of either 0% or 10%. Finally, the criteria consisting of a decrease in size ≥10% measured on three LMs was not statistically superior to the other two with statistical significance (decrease in size ≥0% measured on 3 LMs, and decrease in size ≥10% measured on 5 LMs). The choice to highlight it was based on clinical significance, as the 0% variation threshold has very limited clinical relevance, and measuring three LM is easier, faster and therefore probably more acceptable for clinical practice (and perhaps more reproducible) than measuring five LM.

NETs are characterized by high vascularization, which may imply that tumor efficacy may translate into variation in tumor density. In our study, measuring 10%, 15% or 20% density variations, even combined with size variations, did not enable to cluster responders from

non-responders. However, the relevance of such criteria may be limited to the evaluation of treatments specifically targeting angiogenesis, such as sunitinib. Following its initial description in gastrointestinal tumors, Choi criteria (decrease in size 15% and/or decrease in density ≥10% on portal venous phase) were retrospectively evaluated in patients with advanced pancreatic NETs treated with sunitinib.[27] Several studies reported that the Choi criteria were more sensitive and more precise than RECIST 1.1 in assessing the early response of advanced GEP-NETs treated with sunitinib.[28,29] In the present study, density-based response criteria, including Choi criteria, were not associated with the efficacy of systemic treatments. Should density measurement be irrelevant for evaluating the efficacy of GEP-NET treatments, this would advocate for follow-up using MRI.[30] However, the treatments considered in the present study did not only consist of sunitinib, but also included chemotherapy and somatostatin analogs. The limited number of patients treated with sunitinib precluded subgroup analyses from being performed. Hence, while density-based criteria remain of great interest in the field of NETs, their relevance may be limited to antiangiogenic therapies.[24]

Our study presents some potential biases inherent to any retrospective research. CT analyses were not performed centrally, but all participating institutions are high-volume, NET-dedicated ENETS centers of excellence, with experienced imaging physicians. In addition, all imaging analyses were performed according to a common strict protocol. Interobserver variability could not be evaluated. Nevertheless, CT-scan measurements are usually reproducible in GEP-NET imaging.[14,22,30] Also, patient inclusions were limited by the lack of extensive imaging databases in many expert institutions. To optimize the homogeneity of our series, many patients met exclusion criteria related to the delays between baseline CT-scan, treatment initiation and first evaluation CT-scan, underlining possible differences in routine management across institutions. The inclusion of non-consecutive patients could have induced a selection bias, with however limited consequences on the evaluation of response criteria. Finally, many solicited centers could not participate in this study because follow-up of patients with GEP-NETs is not performed routinely with CT but MRI. However, MRI stands as a very promising technique for evaluating GEP-NET response to treatments, because it may allow more reproducible measurement than CT, does not necessarily require contrast injection, and is a non-radiating technique.[30]

In our study, imaging measurement concerned LMs but not extrahepatic targets, which did not allow progression-free survival to be considered as an endpoint. Instead, we used TTF as the primary endpoint, as it allows global progression to be taken into account in a retrospective setting. While treatment failure was evaluated during the actual management of each patient in each center, it reflects real-life management in expert centers which follow similar guidelines and standards of practice. Finally, we excluded treatments using PRRT due to the limited number of patients, which might limit the clinical implications of this study given that PRRT is now an established second-line treatment option for advanced NETs. However, the treatments included in this cohort (somatostatin analogs, targeted therapies and chemotherapy) remain strong standards today; it is therefore still relevant to identify criteria enabling measurement of their efficacy. Future trials aiming to validate the criteria proposed in this pilot study must focus on patients treated with PRRT.

In conclusion, a decrease ≥10% in the size of three LMs might be a more clinically relevant alternative response criterion than RECIST 1.1 for the early evaluation of treatment efficacy in patients with advanced NETs. Conversely, criteria based on density variation were not predictive of treatment efficacy. This new response criterion should be implemented in clinical trials, in order to validate prospectively the results of this pilot study in a larger population of patients.

AUTHOR CONTRIBUTIONS

Louis de Mestier: Conceptualization; data curation; formal analysis; methodology; supervision; validation; writing – original draft; writing – review and editing. **Matthieu Resche-Rigon:** Data curation; formal analysis; methodology; software; supervision; validation; writing – original draft; writing – review and editing. **Clarisse Dromain:** Data curation; resources; writing – original draft; writing – review and editing. **Angela Lamarca:** Data curation; resources; validation; writing – original draft. **Anna La Salvia:** Data curation; resources; validation; writing – original draft; writing – review and editing. **Lesley de Baker:** Data curation; validation; writing – original draft; writing – review and editing. **Uli Fehrenbach:** Data curation; validation; writing – original draft; writing – review and editing. **Sara Pusceddu:** Data curation; validation; writing – original draft; writing – review and editing. **Annamaria Colao:** Data curation; validation; writing – original draft; writing – review and editing. **Ivan Borbath:** Data curation; validation; writing – original draft; writing – review and editing. **Robbert de Haas:** Data curation; validation; writing – original draft; writing – review and editing. **Maria Rinzivillo:** Data curation; validation; writing – original draft; writing – review and editing. **Alessandro Zerbi:** Data curation; validation; writing – original draft; writing – review and editing. **Luigi Funicelli:** Data curation; validation; writing – original draft; writing – review and editing. **Wouter W. de Herder:** Data curation; validation; writing – original draft; writing – review and editing. **Andreas Selberherr:** Data curation; validation; writing – original draft; writing – review and editing. **Anna Dorothea Wagner:** Data curation; validation; writing – original draft; writing – review and editing. **Prakash Manoharan:** Data curation; validation; writing – original draft; writing – review and editing. **Andrea De Cima:** Data curation; validation; writing – original draft; writing – review and editing. **Willem Lybaert:** Data curation; validation; writing – original draft; writing – review and editing. **Henning Jann:** Data curation; validation; writing – original draft; writing – review and editing. **Natalie Prinzi:** Data curation; validation; writing – original draft; writing – review and editing. **Antongiulio Faggiano:** Data curation; validation; writing – original draft; writing – review and editing. **Laurence Annet:** Data curation; validation; writing – original draft; writing – review and editing. **Annemiek Walenkamp:** Data curation; supervision; writing – original draft; writing – review and editing. **Francesco Panzuto:** Data curation; validation; writing – original draft; writing – review and editing. **Vittorio Pedicini:** Data curation; validation; writing – original draft; writing – review and editing. **Maria Giovanna Pitoni:** Data curation; validation; writing – original draft; writing – review and editing.

Alexander Siebenhuener: Data curation; validation; writing – original draft; writing – review and editing. **Marius E. Mayerhoefer:** Data curation; validation; writing – original draft; writing – review and editing. **Philippe Ruszniewski:** Conceptualization; data curation; formal analysis; methodology; supervision; validation; writing – original draft; writing – review and editing. **Marie-Pierre Vullierme:** Conceptualization; data curation; formal analysis; methodology; project administration; supervision; validation; writing – original draft; writing – review and editing.

AFFILIATIONS

[1]Department of Pancreatology and Digestive Oncology, Université Paris-Cité, INSERM U1149, Beaujon University Hospital, Clichy, France
[2]Department of Epidemiology and Biostatistics, Université Paris-Cité, Saint-Louis Hospital, Paris, France
[3]Department of Radiology, Lausanne University Hospital (CHUV) and University of Lausanne (UNIL), Lausanne, Switzerland
[4]Department of Medical Oncology, The Christie Hospital, Manchester, UK
[5]Department of Medical Oncology, Hospital Universitario 12 de Octubre, Madrid, Spain
[6]Department of Radiology, Universitair Ziekenhuis Antwerpen, Edegem, Belgium
[7]Department of Radiology, Charité – Universitätsmedizin Berlin, Berlin, Germany
[8]Department of Medical Oncology, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy
[9]Endocrinology Unit, Department of Clinical Medicine and Surgery, Università Federico II di Napoli, Naples, Italy
[10]Endocrinology Unit, Department of Clinical and Molecular Medicine, Sapienza University of Rome, Sant'Andrea Hospital, ENETS Center of Excellence, Rome, Italy
[11]Department of Hepatology and Gastroenterology, University Hospital St Luc/UCLouvain, Woluwe, Belgium
[12]Department of Radiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands
[13]Digestive Disease Unit, Sant'Andrea University Hospital, ENETS Center of Excellence, Rome, Italy
[14]Pancreatic Surgery, Humanitas Clinical and Research Center, Rozzano-, Milano, Italy
[15]Division of Radiology, IEO, European Institute of Oncology IRCCS, Milan, Italy
[16]Department of Internal Medicine, Erasmus MC and Erasmus MC Cancer Institute, Rotterdam, The Netherlands
[17]Division of General Surgery, Department of Surgery, Medical University, Vienna, Austria
[18]Department of General and Visceral Surgery, Evangelisches Krankenhaus Wien, Vienna, Austria
[19]Department of Medical Oncology, Lausanne University Hospital (CHUV) and University of Lausanne (UNIL), Lausanne, Switzerland
[20]Department of Radiology and Nuclear Medicine, The Christie, Manchester, UK
[21]Department of Radiology, Hospital Universitario 12 de Octubre, Madrid, Spain
[22]Department of Medical Oncology, Universitair Ziekenhuis Antwerpen, Edegem, Belgium
[23]Department of Hepatology and Gastroenterology, Charité-University, Charité-Universitätsmedizin, Berlin, Germany
[24]Department of Radiology, Cliniques Universitaires Saint-Luc/UCLouvain, Brussels, Belgium
[25]Department of Medical Oncology, University Medical Centre Groningen, University of Groningen, Groningen, The Netherlands
[26]Department of Medical-Surgical Sciences and Translational Medicine, Sapienza University of Rome, ENETS Center of Excellence, Rome, Italy
[27]Department of Radiology, Humanitas Clinical and Research Center, Rozzano-Milano, Italy
[28]Post-graduate School in Radiodiagnostics, University of Milan, Milan, Italy
[29]Department of Gastroenterology and Hepatology, University of Zurich, University Hospital Zurich, Zurich, Switzerland
[30]Department of Biomedical Imaging and Image-guided Therapy, Medical University of Vienna, Vienna, Austria
[31]Department of Radiology, Memorial Sloan Kettering Cancer Center, New York, New York, USA
[32]Department of Radiology, Université Paris-Cité, Beaujon Hospital, Clichy, France

CONFLICT OF INTEREST STATEMENT

LM: AAA, Ipsen, Keocyt, SIRTex. MRR, ALS, BOB, AC, RH, MR, AZ, LF, PM, WL, HJ, AF, LA, AW, FP, VP, GP, PR: none. CD: Received honoraria for Advisory Board Meeting from Ipsen. AL: received travel and educational support from Ipsen, Pfizer, Bayer, AAA, SirtEx, Novartis, Mylan and Delcath, speaker honoraria from Merck, Pfizer, Ipsen, Incyte, AAA, QED, Servier, Astra Zeneca and EISAI; advisory and consultancy honoraria from EISAI, Nutricia Ipsen, QED, Roche, Servier, Boston Scientific, Albireo Pharma, AstraZeneca, Boehringer Ingelheim, GENFIT and TransThera Biosciences, member of the Knowledge Network and NETConnect Initiatives funded by Ipsen. UF: GE, Bayer, Siemens, Ipsen, ESGAR, Asahi Intecc. SP: received honoraria from Novartis, Ipsen, Pfizer, MerckSerono, and Advanced Accelerator Applications – AAA; received institutional research grant by Ipsen, Pfizer; IB: received travel and educational support from Ipsen, Servier, Roche, speaker honoraria from Servier, Ipsen, Astra Zeneca and EISAI; advisory and consultancy honoraria from EISAI, QED, Roche, Servier, AstraZeneca. WWH: Received honoraria from Novartis and Ipsen, and participated in Advisory Boards of Novartis and Ipsen. ARS: received consulting and speaker honoraria from Amgen, AstraZeneca, AAA, Bayer, BMS, IPSEN, Lilly, Merck, MSD, Novartis, Pfizer, Roche, Sanofi and Servier for work performed outside the current study. ADW: received consulting fees and speaker honoraria from Merck, Lily, Pierre-Fabre Pharma, Hofmann-LaRoche, Sanofi, Daiichi-Sankyo, Dragonfly Therapeutics, Servier, Bristol-Myers-Squibb, Astellas. She is coordinating investigator of EORTC trial 1203, which is supported by an educational grant from Roche to

EORTC. NP: Ipsen, Pfizer, Novartis, Italfarmaco, AAA, Merk. MEM: Siemens, GE, BMS. MEM was funded in part through the NIH/NCI Cancer Center Support Grant P30 CA008748. MPV: received travel and educational support from Ipsen, Received honoraria from Guerbet, International.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/jne.13311.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from Louis de Mestier, Matthieu Resche-Rigon and Marie-Pierre Vullierme upon reasonable request.

ORCID

Clarisse Dromain https://orcid.org/0000-0002-2999-1958
Uli Fehrenbach https://orcid.org/0000-0003-3622-3268
Ivan Borbath https://orcid.org/0000-0001-7598-214X
Wouter W. de Herder https://orcid.org/0000-0003-1463-5165
Andreas Selberherr https://orcid.org/0000-0002-6545-7023
Marie-Pierre Vullierme https://orcid.org/0000-0002-3667-8943

REFERENCES

1. Dasari A, Shen C, Halperin D, et al. Trends in the incidence, prevalence, and survival outcomes in patients with neuroendocrine tumors in the United States. *JAMA Oncol*. 2017;3:1335-1342. doi:10.1001/jamaoncol.2017.0589
2. Borbath I, Garcia-Carbonero R, Bikmukhametov D, et al. The European Neuroendocrine Tumour Society registry, a tool to assess the prognosis of neuroendocrine neoplasms. *Eur J Cancer*. 2022;168:80-90. doi:10.1016/j.ejca.2022.03.007
3. Pavel M, Öberg K, Falconi M, et al. Gastroenteropancreatic neuroendocrine neoplasms: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2020;31:844-860. doi:10.1016/j.annonc.2020.03.304
4. de Mestier L, Lepage C, Baudin E, et al. Digestive neuroendocrine neoplasms (NEN): French intergroup clinical practice guidelines for diagnosis, treatment and follow-up (SNFGE, GTE, RENATEN, TENPATH, FFCD, GERCOR, UNICANCER, SFCD, SFED, SFRO, SFR). *Dig Liver Dis*. 2020;52:473-492. doi:10.1016/j.dld.2020.02.011
5. Rinke A, Müller H-H, Schade-Brittinger C, et al. Placebo-controlled, double-blind, prospective, randomized study on the effect of octreotide LAR in the control of tumor growth in patients with metastatic neuroendocrine midgut tumors: a report from the PROMID study group. *JCO*. 2009;27:4656-4663. doi:10.1200/JCO.2009.22.8510
6. Caplin ME, Pavel M, Cwikła JB, et al. Lanreotide in metastatic enteropancreatic neuroendocrine tumors. *N Engl J Med*. 2014;371:224-233. doi:10.1056/NEJMoa1316158
7. Yao JC, Fazio N, Singh S, et al. Everolimus for the treatment of advanced, non-functional neuroendocrine tumours of the lung or gastrointestinal tract (RADIANT-4): a randomised, placebo-controlled, phase 3 study. *Lancet*. 2016;387:968-977. doi:10.1016/S0140-6736(15)00817-X
8. Yao JC, Shah MH, Ito T, et al. Everolimus for advanced pancreatic neuroendocrine tumors. *N Engl J Med*. 2011;364:514-523. doi:10.1056/NEJMoa1009290
9. Raymond E, Dahan L, Raoul J-L, et al. Sunitinib malate for the treatment of pancreatic neuroendocrine tumors. *N Engl J Med*. 2011;364:501-513. doi:10.1056/NEJMoa1003825
10. Strosberg J, El-Haddad G, Wolin E, et al. Phase 3 trial of 177Lu-Dotatate for midgut neuroendocrine tumors. *N Engl J Med*. 2017;376:125-135. doi:10.1056/NEJMoa1607427
11. Moertel CG, Lefkopoulo M, Lipsitz S, Hahn RG, Klaassen D. Streptozocin-doxorubicin, streptozocin-fluorouracil or chlorozotocin in the treatment of advanced islet-cell carcinoma. *N Engl J Med*. 1992;326:519-523. doi:10.1056/NEJM199202203260804
12. Kunz PL, Catalano PJ, Nimeiri H, et al. A randomized study of temozolomide or temozolomide and capecitabine in patients with advanced pancreatic neuroendocrine tumors: a trial of the ECOG-ACRIN Cancer Research Group (E2211). *J Clin Oncol*. 2018;36:abstract 4004. doi:10.1200/JCO.2018.36.15_suppl.4004
13. 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. doi:10.1016/j.ejca.2008.10.026
14. Sundin A, Arnold R, Baudin E, et al. ENETS Consensus Guidelines for the Standards of care in neuroendocrine tumors: radiological, nuclear medicine and hybrid imaging. *Neuroendocrinology*. 2017;105:212-244. doi:10.1159/000471879
15. de Mestier L, Dromain C, d'Assignies G, et al. Evaluating digestive neuroendocrine tumor progression and therapeutic responses in the era of targeted therapies: state of the art. *Endocr Relat Cancer*. 2014;21:R105-R120. doi:10.1530/ERC-13-0365
16. Merino-Casabiel X, Aller J, Arbizu J, et al. Consensus document on the progression and treatment response criteria in gastroenteropancreatic neuroendocrine tumors. *Clin Transl Oncol*. 2018;20:1522-1528. doi:10.1007/s12094-018-1881-9
17. Thiis-Evensen E, Poole AC, Nguyen H-TT, Sponheim J. Achieving objective response in treatment of non-resectable neuroendocrine tumors does not predict longer time to progression compared to achieving stable disease. *BMC Cancer*. 2020;20:466. doi:10.1186/s12885-020-06963-6
18. Choi H, Charnsangavej C, Faria SC, et al. Correlation of computed tomography and positron emission tomography in patients with metastatic gastrointestinal stromal tumor treated at a single institution with imatinib mesylate: proposal of new computed tomography response criteria. *J Clin Oncol*. 2007;25:1753-1759. doi:10.1200/JCO.2006.07.3049
19. Lencioni R, Llovet JM. Modified RECIST (mRECIST) assessment for hepatocellular carcinoma. *Semin Liver Dis*. 2010;30:52-60. doi:10.1055/s-0030-1247132
20. WHO Classification of Tumours. *Digestive System Tumours*. Vol. 1. 5th ed. IARC; 2019.
21. Perren A, Couvelard A, Scoazec J-Y, et al. ENETS consensus guidelines for the standards of care in neuroendocrine tumors: pathology: diagnosis and prognostic stratification. *Neuroendocrinology*. 2017;105:196-200. doi:10.1159/000457956
22. Zappa M, Hentic O, Vullierme M-P, et al. Is visual radiological evaluation of liver tumour burden in patients with neuroendocrine tumours reproducible? *Endocr Connect*. 2017;6:33-38. doi:10.1530/EC-16-0092
23. Therasse P, Arbuck SG, Eisenhauer EA, et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst*. 2000;92:205-216.
24. Capdevila J, Grande E, García-Carbonero R, et al. Position statement on the diagnosis, treatment, and response evaluation to systemic therapies of advanced neuroendocrine tumors, with a special focus on radioligand therapy. *Oncologist*. 2022;27:e328-e339. doi:10.1093/oncolo/oyab041

25. Lamarca A, Barriuso J, Kulke M, et al. Determination of an optimal response cut-off able to predict progression-free survival in patients with well-differentiated advanced pancreatic neuroendocrine tumours treated with sunitinib: an alternative to the current RECIST-defined response. *Br J Cancer*. 2018;118:181-188. doi:10.1038/bjc.2017.402
26. Lamarca A, Ronot M, Moalla S, et al. Tumor growth rate as a validated early radiological biomarker able to reflect treatment-induced changes in neuroendocrine tumors: the GREPONET-2 study. *Clin Cancer Res*. 2019;25:6692-6699. doi:10.1158/1078-0432.CCR-19-0963
27. Faivre S, Ronot M, Dreyer C, et al. Imaging response in neuroendocrine tumors treated with targeted therapies: the experience of sunitinib. *Target Oncol*. 2012;7:127-133. doi:10.1007/s11523-012-0216-y
28. Luo Y, Chen J, Huang K, et al. Early evaluation of sunitinib for the treatment of advanced gastroenteropancreatic neuroendocrine neoplasms via CT imaging: RECIST 1.1 or Choi criteria? *BMC Cancer*. 2017;17:154. doi:10.1186/s12885-017-3150-7
29. Solis-Hernandez MP, Fernandez del Valle A, Carmona-Bayonas A, et al. Evaluating radiological response in pancreatic neuroendocrine tumours treated with sunitinib: comparison of Choi versus RECIST criteria (CRIPNET_ GETNE1504 study). *Br J Cancer*. 2019;121:537-544. doi:10.1038/s41416-019-0558-7
30. Moalla S, Arfi Rouche J, Foulon S, et al. Are we reproducible in measurement of NET liver metastasis? *Dig Liver Dis*. 2017;49:1121-1127. doi:10.1016/j.dld.2017.05.015

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: de Mestier L, Resche-Rigon M, Dromain C, et al. Proposal of early CT morphological criteria for response of liver metastases to systemic treatments in gastroenteropancreatic neuroendocrine tumors: Alternatives to RECIST. *J Neuroendocrinol*. 2023;35(6):e13311. doi:10.1111/jne.13311