**Radiological response heterogeneity is of prognostic significance in metastatic renal cell carcinoma treated with vascular endothelial growth factor-targeted therapy**

Peter E Halla, Scott TC Shepherdb,c, Janet Brownd,e, James Larkinc, Robert Jonesf, Christy Ralphd, Robert Hawkinsg, Simon Chowdhuryh, Ekaterini Boletib, Amit Bahli, Kate Fifej, Andrew Webbk, Simon J Crabbl, Thomas Geldartm, Robert Hilln, Joanna Dunlopn, Duncan McLareno, Charlotte Ackermana, Akhila Wimalasinghama, Luis Beltrana, Paul Nathanp, Thomas Powlesa,b,†

a Barts Cancer Institute, CRUK Experimental Cancer Medicine Centre, London, UK.

b Dept. of Oncology, Royal Free NHS Foundation Trust, London, UK.

c Dept. of Medical Oncology, Royal Marsden Hospital, London, UK.

d Dept. of Medical Oncology, Leeds Teaching Hospitals NHS Trust, Leeds, UK

e Academic Unit of Clinical Oncology, Dept. of Oncology and Metabolism, University of Sheffield, Sheffield, UK.

f Beatson Cancer Centre, University of Glasgow, Glasgow, UK.

g Dept. of Medical Oncology, Christie Hospital, Manchester, UK.

h Dept. of Oncology, Guys and St Thomas’ NHS Foundation Trust, London, UK.

i Dept. of Oncology, University Hospital Bristol NHS Foundation trust, Bristol, UK.

j Dept. of Oncology, Cambridge University Hospitals, Cambridge, UK.

k Dept. of Oncology, Brighton and Sussex University Hospital Trust, Brighton, UK.

l Cancer Sciences Unit, University of Southampton, Southampton, UK.

m Dept. of Oncology, Royal Bournemouth Hospital, Bournemouth, UK.

n Scottish Clinical Trials Research Unit (SCTRU), NHS National Services Scotland, Edinburgh, UK.

o Edinburgh Cancer Centre, Western General Hospital, Edinburgh, UK.

p Dept. of Oncology, Mount Vernon Cancer Centre, Northwood, UK.

†**Corresponding author:**

Professor Thomas Powles

Renal Lead for London Cancer

Department of Oncology

Royal Free NHS Foundation Trust

London NW3 2QG, UK

Tel: +44 (0)20 77940500 extension 33667 or direct line +44 (0)20 7472 6778

Fax: +44 (0)20 7794 3341

Email: thomas.powles@bartshealth.nhs.uk

**Word count (abstract): 298**

**Word count (text including abstract): 2,792**

**ABSTRACT**

**Background:** Response Evaluation Criteria in Solid Tumours (RECIST) is widely used to assess tumour response but is limited by not considering disease site or radiological heterogeneity.

**Objective:** To determine if radiological heterogeneity or disease site have prognostic significance in patients with metastatic clear cell renal cell carcinoma (ccRCC).

**Design, Setting, and Participants:** A retrospective analysis was conducted of a second-line phase II study evaluating VEGF-targeted therapy ± Src inhibition in metastatic ccRCC (NCT00942877). 138 patients with 458 baseline lesions were assessed.

**Outcome Measurements and Statistical Analysis:** Radiological heterogeneity at week 8 was assessed within individual patients with ≥2 lesions to predict overall survival (OS) using Kaplan-Meier method and Cox regression. We defined a high heterogeneous response as occurring when ≥1 lesion underwent a ≥10% reduction *and* ≥1 lesion underwent a ≥10% increase in size. Disease progression was defined by RECIST 1.1 criteria.

**Results and Limitation:** In patients with a complete/partial response or stable disease by RECIST 1.1 and ≥2 lesions at week 8, those with a high heterogeneous response had a shorter OS compared to those with a homogeneous response (hazard ratio [HR] 2.01 (95% confidence interval [CI]: 1.39-2.92; *P*<0.001). Response by disease site at week 8 did not affect OS. At disease progression, ≥1 new lesion was associated with worse survival compared to >20% increase in sum of target lesion diameters only (HR 2.12; 95% CI: 1.43-3.14; *P*<0.001). Limitations include retrospective study design.

**Conclusions:** Radiological heterogeneity and the development of new lesions may predict survival in metastatic ccRCC. Further prospective studies are required to validate these findings.

**Patient summary:** We looked at individual metastases in patients with kidney cancer and showed that a variable response to treatment and the appearance of new metastases may be associated with worse survival. Further studies are required to confirm these findings.

**KEY WORDS**

Heterogeneity; prognostic factor; radiological response; renal cell carcinoma; vascular endothelial growth factor.

**INTRODUCTION**

Inhibition of vascular endothelial growth factor (VEGF) signalling, usually by means of small-molecule tyrosine kinase inhibitors (TKIs), is the current mainstay of metastatic clear-cell renal cell carcinoma (ccRCC) therapy in both the first and second-line settings [1]. However, there is a wide variation in treatment responses by patients. Several prognostic scoring systems have been developed to identify poor and favourable risk patients [2, 3]. These are determined at baseline and are based around a combination of time to treatment, performance status and blood parameters [2, 3]. Response Evaluation Criteria in Solid Tumours (RECIST) response rates and disease progression have been used as surrogate markers of activity in clinical trials [4]. However, RECIST is limited as it overlooks details of dynamic changes by amalgamating total tumour burden into a single numerical entity. Confidence in RECIST as accurate surrogate marker of outcome is also questionable, partly due to variable responses within individual patients, also known as intra-patient heterogeneity. For these reasons, clinicians often continue treatment past disease progression. Therefore, more accurate tools for predicting outcome are required. We hypothesised that following individual lesion responses would better characterise clinical benefit. We therefore examined individual lesions in patients with metastatic ccRCC participating in a VEGF-targeted therapy clinical trial to address this.

**PATIENTS AND METHODS**

*Study population*

Prospectively collected data from the double-blind, randomised, phase II COSAK trial (ClinicalTrials.gov NCT00942877) were used in this retrospective *post hoc* analysis. One hundred and thirty-eight patients with metastatic ccRCC who had progressed after at least one line of VEGF targeted therapy were randomised to either cediranib (a VEGF TKI) alone (*N*=69) or in combination with the Src inhibitor, saracatinib (*N*=69). Exclusion criteria included untreated brain metastases, uncontrolled hypertension and concurrent malignancies. The two arms were well-matched for patient characteristics. No significant difference was seen in progression-free survival (PFS) or overall survival (OS). The results of this trial have been published elsewhere [5]. Given that no difference was seen between the two treatment arms, the data were combined for this analysis.

*Imaging and image analysis*

Computed tomography (CT) scans were undertaken every eight weeks using standard protocols and patient response assessed by RECIST 1.1 [4]. The radiological analyses were undertaken by staff blinded to the outcome data, but no central review occurred. Baseline, week 8 and disease progression were the time points examined. Individual lesion responses (percentage change from baseline) for each patient were also determined at week 8. RECIST 1.1 criteria were used to categorise each lesion response.

*Radiological heterogeneity*

Radiological heterogeneity (RH) was assessed at week 8 in patients with ≥2 lesions using criteria developed by van Kessel and colleagues in patients with colorectal liver metastases (Suppl. Fig. 1, [6]). Van Kessel and colleagues used the terms ‘homogeneous’, ‘mixed’ and ‘true mixed’ response, but these may be misleading as ‘true’ implies a validated comparison with a gold standard. We therefore have used the category terms ‘homogeneous’, ‘low heterogeneous’ and ‘high heterogeneous’ response instead.

Briefly, the percentage change in each lesion was determined and the maximum difference calculated. A homogeneous response indicated that all the lesions for a patient had changed in the same direction with <30% difference between highest and lowest change. A low heterogeneous response indicated that all lesions changed in same direction, but that there was a ≥30% difference between the highest and lowest. For the homogeneous and low heterogeneous response categories, small changes (-10% to +10%) could be re-assigned to count as a change in the same direction. A high heterogeneous response indicated that at least one lesion underwent a ≥10% reduction *and* at least one other lesion underwent a ≥10% increase.

*Statistical analysis*

The primary outcome for this study was OS. Kaplan-Meier method was used to assess OS and groups compared using the log-rank test. Uni- and multivariate analyses were undertaken using Cox regression to calculate Hazard Ratios (HR) and adjust for other prognostic variables (age, gender, Eastern Cooperative Oncology Group [ECOG] performance status, Memorial Sloan Kettering Cancer Center [MSKCC] risk group [2], nephrectomy status). Pearson’s Chi-Square test was used to assess differences in RH between two groups. All statistical analyses were conducted using Statistical Package for the Social Sciences (SPSS, version 23). A *P* value of <0.05 was considered significant.

**RESULTS**

*Patients*

All 138 patients from the COSAK trial were evaluated (Table 1). One hundred and eight (78%) were male. The median age was 60 years. Ninety-six percent of patients had received only one previous VEGF-targeted therapy whereas the remainder had received two previous lines of VEGF therapy. Eighty-one percent of patients had previously received sunitinib. Overall, 14%, 72% and 14% were in the good, intermediate and poor MSKCC risk group at randomisation [2]. Median PFS and OS for the whole group were 4.1 months (95% confidence interval (CI): 3.1-5.1 months) and 12.0 months (95% CI: 8.5-15.6 months), respectively. No significant difference between the treatment arms was observed with regards to both baseline characteristics and treatment response (*P*>0.05). Therefore, the two treatment arms were merged for this analysis.

*Baseline site of disease*

At baseline, 458 individual lesions from 138 patients were available for analysis. The median number of lesions per patient was 3 (range 1-5). A breakdown of the lesion sites was as follows: lymph nodes 138 (30%); lung 112 (24%); liver 42 (9%); bone 27 (6%); other 139 (30%). Twenty-seven patients had ≥ 1 liver metastasis (20%) and 18 (13%) had ≥ 1 bone metastasis. Two patients (1.4%) had both a liver and bone metastasis. The presence of a liver or bone metastasis was not predictive of PFS (HR 0.95; 95% CI: 0.66-1.38; *P*=0.80) or OS (HR 1.34; 95% CI: 0.91-1.97; *P*=0.14).

*First follow-up CT scan (week 8)*

The first follow-up CT scan occurred at week 8. One hundred and thirteen patients (82% of baseline) had week 8 data for analysis encompassing 369 of the baseline lesions (81%; lymph nodes 103 [28% of the 369], Lung 93 [25%], liver 30 [8%], bone 26 [7%], other 117 [32%]). Reasons for the reduced patient numbers at week 8 included death and drug toxicity.

*Individual lesion responses at week 8*

Assessment of the individual lesion responses at week 8 by RECIST criteria showed one complete response (0.3%), 49 partial responses (13%), 276 (75%) were classified as stable and 43 (12%) lesions progressed (Suppl. Table 1A). Lesion site responses of CR/PR (combined as only one lesion had a CR), SD or PD were not prognostic for OS (Suppl. Table 1B).

*Overall patient responses at week 8*

When overall patient responses were analysed by RECIST at week 8, no patients had a CR, 8 (7.1%) had a PR, 80 (70.8%) had SD and 25 (22.1%) had PD. As expected, PD at week 8 was associated with worse OS with a median of 3.9 months (95% CI: 1.0-6.8) compared to 12.1 months (95% CI: 9.7-14.5; HR 1.61; 95% CI: 1.07-2.43; *P*=0.02) for patients with a PR and 13.9 months (95% CI: 12.2-15.6; HR 3.21; 95% CI: 2.10-4.93; *P*<0.001) for patients with SD. No statistical difference was seen between the PR and SD groups (HR 0.82; 95% CI: 0.37-1.79; *P*=0.61).

*Radiological heterogeneity at week 8*

Given that no difference in outcome was seen between the RECIST-defined PR and SD groups at week 8, we examined whether OS in this subpopulation could be further characterised by RH. Of the 113 patients with individual lesion data available at week 8, 104 (75% of the initial 138 patients) had >1 lesion and therefore could be assessed for heterogeneity. Of these 104 patients, 81 (59% of the initial 138 patients), had PR (*N*=7) or SD (*N*=74) by RECIST at week 8 and were included in the heterogeneity analysis. The remaining 23 patients had PD by RECIST criteria and were not included.

Figure 1 demonstrates the frequency of different lesion responses by RECIST category for PR and SD patients combined. Radiological heterogeneity was commonly seen, with 34 patients (42%) having ≥2 RECIST categories amongst their lesion responses at week 8. However, heterogeneity by number of RECIST categories (1 versus ≥2) was not associated with improved OS (HR 1.40; 95% CI: 0.84-2.32; *P*=0.19).

Radiological heterogeneity was assessed using criteria developed for colorectal liver metastases in the RECIST-defined PR and SD populations (Suppl. Fig. 1; [6]). Forty nine patients (60%) had a homogeneous response, 20 (25%) had a low heterogeneous response and 12 (15%) had a high heterogeneous response by RH criteria. For OS from week 8, the times were 16.9 months (Fig. 2; 95% CI: 11.1-22.7), 12.8 months (95% CI: 11.3-14.3) and 7.3 months (95% CI: 5.4-9.2) for the homogeneous, low heterogeneous and high heterogeneous response categories, respectively. Hazard ratios were: Homogeneous *vs* low heterogeneous 1.41 (95% CI: 0.78-2.55; *P*=0.26); Homogeneous *vs* high heterogeneous 2.01 (95% CI: 1.39-2.92; *P*<0.001); low heterogeneous *vs* high heterogeneous 2.58 (95% CI: 1.12-5.91; *P*=0.02).

We hypothesised that patients with smaller, more numerous lesions may demonstrate increased RH and that therefore these were confounding our results. Of the 81 patients in the RH analysis, 28 (35%) had two target lesions and 53 (65%) had ≥3 lesions. The number of target lesions (2 *vs* ≥3) was not prognostic for OS (HR 0.66; 95% CI: 0.39-1.12; *P*=0.13). The median sum of target lesion diameters at week 8 was 92 mm (range 20-334). A sum below the median was associated with improved OS (HR 0.45; 95% CI: 0.27-0.74; *P*=0.002), but RH was not significantly different between the two groups (Suppl. Fig. 2; *P*=0.17). However, in a multivariate Cox regression including RH, sum of lesion diameters, number of lesions alongside the other variables, only RH, sum of lesion diameters and MSKCC score were independent prognostic factors for OS (Table 2).

Radiological heterogeneity was not prognostic for OS in patients with PD at week 8, although numbers were small (HR 0.76, 95% CI: 0.31-1.83; *P*=0.54; *N*=23).

*New lesions at disease progression predict worse survival*

One hundred and twenty one patients (88% of the initial 138) had data at disease progression. Of these, 64 (53%) had no new sites of disease and 57 (47%) had ≥1 new site. Lung was the commonest site of for a new lesion (23 patients, 41%) with liver and ‘other’ being the next commonest sites (16 patients each, 29%). This was followed by bone (13 patients, 21%), lymph node (8 patients, 14%) and brain (2 patients, 4%). The new site was unknown for one patient. Median survival was significantly shorter in patients with ≥1 new site of disease compared to none at disease progression (Fig. 3; 3.7 months [95% CI: 2.1-5.2] versus 9.9 months [95% CI: 7.5-12.2]; HR 2.12; 95% CI: 1.43-3.14; *P*<0.001). In patients with ≥1 new disease site, 32 patients (56%) had a <20% increase in the sum of lesion diameters at disease progression, 21% had ≥20% increase and 23% had missing data. No significant difference in survival was seen between the groups suggesting new sites rather than general progression in all sites was associated with poor outcome (HR 0.87; 95% CI: 0.42-1.79; *P*=0.66). The site of the new lesion was not predictive for survival (Suppl. Table 2).

**DISCUSSION**

This study examined radiological prognostic factors at baseline, first follow-up scan (week 8) and disease progression in metastatic ccRCC patients receiving second-line VEGF-targeted therapy.

Whilst patients with PD at first follow-up had a worse survival, no significant difference in survival was seen between patients with PR or SD when using RECIST 1.1 criteria. Therefore, alternative radiological prognostic markers were sought for these patients to predict prognosis and thus aid treatment decisions. Forty percent of patients with non-progressive disease at week 8 demonstrated RH, with increased RH associated with worse survival. Intratumoural and inter-metastasis heterogeneity has been shown to exist at a molecular level in RCC where clonal evolution is thought to play a role [7-9]. Similarly, RH has been shown to exist in metastatic ccRCC patients treated with first-line VEGF-targeted therapies at a similar frequency to that seen in this study and is likely to represent different clones [10]. However, no outcome data were analysed. Radiological heterogeneity has also been shown in patients with colorectal liver metastases where increased RH was correlated to a worse OS [6]. We have described a method to assess RH that can be used in the clinic and, in our dataset, had prognostic significance for patients with metastatic ccRCC at their first follow-up scan thereby providing a potential alternative to RECIST for assessing treatment response. This may be beneficial to patients as ineffective treatments can be changed at an earlier timepoint. Radiological heterogeneity was found to be independent of potential confounders, number of target lesions and sum of lesion diameters, but further validation is required. Future studies may also look at the correlation between RH and tumour factors including Fuhrman grade and Von-Hippel-Lindau mutational status.

The development of ≥1 new lesion, rather than the growth of existing lesions, at disease progression was associated with a worse OS. This has previously been described for metastatic RCC patients treated with everolimus [11]. Similar effects have been shown in metastatic breast, colorectal and lung cancer [12, 13]. RECIST does not distinguish between the two types of disease progression, thereby reflecting a further limitation of its use. The development of new sites suggests increased clinical significance and may help decision making in terms of switching therapy.

Baseline site of disease was not a prognostic factor for OS in this study. In addition, treatment response at week 8 by disease site was not prognostic for survival. This is in contrast to previous studies which have shown that bone and liver metastases are adverse independent prognostic factors for OS in metastatic RCC [14-16]. This correlates with findings from patients treated with cytokines where liver and bone metastases have been included as adverse factors in a prognostic model [17]. It is unclear why bone and liver metastases were not prognostic in this study, although low *N* numbers may be one explanation.

There are several limitations of this study. This was a retrospective study that was not powered for the groups analysed and therefore requires validation before definitive conclusions can be reached, ideally with prospective studies. The *N* numbers in this study were small, making it difficult to reject to the null hypothesis. Nonetheless, even with this restriction, we did manage to show significant results. Cediranib is not licensed for use in RCC, having not been developed further due largely to the competitive landscape in metastatic RCC. Its efficacy appears to be in line with other VEGF-targeted therapies tested in the ≥2nd line setting, but further work is required to see if the conclusions from this paper are applicable to other VEGF-targeted therapies in both the first- and second-line settings [5, 18].

**CONCLUSIONS**

In conclusion, we have shown that radiological heterogeneity may have prognostic value at the first follow-up scan and may be a useful factor in determining whether to change treatments. Similarly, the development of new lesions at disease progression is associated with a worse survival than solely an increase in the size of existing lesions. Further prospective validation is required to confirm these findings.

**PATIENT SUMMARY**

We looked at individual metastases in patients with kidney cancer and showed that a variable response to treatment and the appearance of new metastases may be associated with worse survival. Further studies are required to confirm these findings.

**FUNDING AND SPONSORSHIP**

This study was supported by Cancer Research UK and AstraZeneca. It was sponsored by the Common Services Agency (CSA; NHS National Services Scotland).

**CONFLICT OF INTEREST STATEMENT**

PEH has received honoraria from Merck Sharp & Dohme. JB has received honoraria from Amgen, Pfizer and Novartis. RJ has received research funding from AstraZeneca. CR has received sponsorship and honoraria from Pfizer, Novartis, Bristol-Meyers Squibb, Roche, GlaxoSmithKline, Viralytics, Janssen and the British Sarcoma Group. SC has received funding from GlaxoSmithKline and Pfizer for speaking. SJC has received sponsorship from Novartis and Merck, and research funding from AstraZeneca. KF has received honoraria from Roche, Pfizer and Novartis. TP has received honoraria for advisory boards from Novartis, Roche, Pfizer and Bristol-Myers Squibb and received a research grant from AstraZeneca. All remaining authors have declared no conflict of interest.

**REFERENCES**

[1] Escudier B, Porta C, Schmidinger M, Algaba F, Patard JJ, Khoo V, et al. Renal cell carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2014;25 Suppl 3:iii49-56.

[2] Motzer RJ, Bacik J, Murphy BA, Russo P, Mazumdar M. Interferon-alfa as a comparative treatment for clinical trials of new therapies against advanced renal cell carcinoma. J Clin Oncol. 2002;20:289-96.

[3] Heng DY, Xie W, Regan MM, Warren MA, Golshayan AR, Sahi C, et al. Prognostic factors for overall survival in patients with metastatic renal cell carcinoma treated with vascular endothelial growth factor-targeted agents: results from a large, multicenter study. J Clin Oncol. 2009;27:5794-9.

[4] Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45:228-47.

[5] Powles T, Brown J, Larkin J, Jones R, Ralph C, Hawkins R, et al. A randomized, double-blind phase II study evaluating cediranib versus cediranib and saracatinib in patients with relapsed metastatic clear-cell renal cancer (COSAK). Ann Oncol. 2016;27:880-6.

[6] van Kessel CS, Samim M, Koopman M, van den Bosch MA, Borel Rinkes IH, Punt CJ, et al. Radiological heterogeneity in response to chemotherapy is associated with poor survival in patients with colorectal liver metastases. Eur J Cancer. 2013;49:2486-93.

[7] Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, Gronroos E, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med. 2012;366:883-92.

[8] Martinez P, Birkbak NJ, Gerlinger M, McGranahan N, Burrell RA, Rowan AJ, et al. Parallel evolution of tumour subclones mimics diversity between tumours. J Pathol. 2013;230:356-64.

[9] Gerlinger M, Horswell S, Larkin J, Rowan AJ, Salm MP, Varela I, et al. Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. Nat Genet. 2014;46:225-33.

[10] Crusz SM, Tang YZ, Sarker SJ, Prevoo W, Kiyani I, Beltran L, et al. Heterogeneous response and progression patterns reveal phenotypic heterogeneity of tyrosine kinase inhibitor response in metastatic renal cell carcinoma. BMC medicine. 2016;14:185.

[11] Stein A, Bellmunt J, Escudier B, Kim D, Stergiopoulos SG, Mietlowski W, et al. Survival prediction in everolimus-treated patients with metastatic renal cell carcinoma incorporating tumor burden response in the RECORD-1 trial. Eur Urol. 2013;64:994-1002.

[12] Litiere S, de Vries EG, Seymour L, Sargent D, Shankar L, Bogaerts J. The components of progression as explanatory variables for overall survival in the Response Evaluation Criteria in Solid Tumours 1.1 database. Eur J Cancer. 2014;50:1847-53.

[13] Suzuki C, Blomqvist L, Sundin A, Jacobsson H, Bystrom P, Berglund A, et al. The initial change in tumor size predicts response and survival in patients with metastatic colorectal cancer treated with combination chemotherapy. Ann Oncol. 2012;23:948-54.

[14] McKay RR, Kroeger N, Xie W, Lee JL, Knox JJ, Bjarnason GA, et al. Impact of bone and liver metastases on patients with renal cell carcinoma treated with targeted therapy. Eur Urol. 2014;65:577-84.

[15] Beuselinck B, Oudard S, Rixe O, Wolter P, Blesius A, Ayllon J, et al. Negative impact of bone metastasis on outcome in clear-cell renal cell carcinoma treated with sunitinib. Ann Oncol. 2011;22:794-800.

[16] Iacovelli R, Santini D, Rizzo M, Felici A, Santoni M, Verzoni E, et al. Bone metastases affect prognosis but not effectiveness of third-line targeted therapies in patients with metastatic renal cell carcinoma. Can Urol Assoc J. 2015;9:263-7.

[17] Negrier S, Escudier B, Gomez F, Douillard JY, Ravaud A, Chevreau C, et al. Prognostic factors of survival and rapid progression in 782 patients with metastatic renal carcinomas treated by cytokines: a report from the Groupe Francais d'Immunotherapie. Ann Oncol. 2002;13:1460-8.

[18] Rini BI, Escudier B, Tomczak P, Kaprin A, Szczylik C, Hutson TE, et al. Comparative effectiveness of axitinib versus sorafenib in advanced renal cell carcinoma (AXIS): a randomised phase 3 trial. Lancet. 2011;378:1931-9.

**TABLE LEGENDS**

**Table 1: Patients’ characteristics at baseline.** VEGF, vascular endothelial growth factor. MSKCC, Memorial Sloan Kettering Cancer Center.

**Table 2: Multivariate Cox regression analysis of variables affecting overall survival at week 8.** ECOG, Eastern Cooperative Oncology Group; MSKCC, Memorial Sloan Kettering Cancer Center.

**FIGURE LEGENDS**

**Figure 1: Frequencies of individual lesion response categories by RECIST 1.1 at week 8 in patients with non-progressive disease.** Individual lesion responses were assessed according to RECIST 1.1 criteria in patients who had an overall response of either PR or SD at week 8 (no CR by patient). Note, only one lesion had a CR and therefore was combined with the PR group. The types of RECIST category demonstrated by the lesions within a patient were assessed and the number of patient with those categories determined. CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease.

**Figure 2: Radiological heterogeneity in patients with stable disease at week 8 is associated with overall survival.** In patients with stable disease at week 8, radiological heterogeneity (homogeneous response, low heterogeneous response, high heterogeneous response) is prognostic for overall survival: 19.3 months (95% CI: 14.2-24.5), 12.8 months (95% CI: 10.7-14.9) and 7.2 months (95% CI: 5.4-9.2) for the homogeneous response, low heterogeneous response and high heterogeneous response categories, respectively. Hazard ratios were as follows: Homogeneous *vs* Low heterogeneous 1.41 (95% CI: 0.78-2.55; *P*=0.26); Homogeneous *vs* High heterogeneous 2.01 (95% CI: 1.39-2.92; *P*<0.001); Low heterogeneous *vs* High heterogeneous 2.58 (95% CI: 1.12-5.91; *P*=0.02).

**Figure 3: One or more new lesion at disease progression is associated with worse overall survival.** %). Median survival was significantly shorter in patients with ≥1 new site of disease compared to none at disease progression (3.7 months [95% CI: 2.1-5.2] versus 9.9 months [95% CI: 7.5-12.2]; HR 2.12; 95% CI: 1.43-3.14; *P*<0.001).

**SUPPLEMENTARY TABLE/FIGURE LEGENDS**

**Supplementary Table 1: Lesion response by site at first follow-up scan (week 8).** Individual lesion responses were assessed by RECIST 1.1 criteria (A). Only one lesion had a CR at week 8 and was therefore combined with the PR category. Hazard ratios for overall survival (OS) were analysed by site and RECIST 1.1 response in individual lesions (B). None were predictive for OS (*P*>0.05). CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease; N/A, not applicable due to low *N* numbers; ( ) indicate 95% confidence interval.

**Supplementary Table 2: Site of new lesion at disease progression does not predict survival.** Hazard ratios for the site of the new lesion at disease progression compared to all other sites. N/A, not applicable due to low *N* numbers.

**Supplementary Figure 1: Methods used to assess radiological response heterogeneity.** Radiological response heterogeneity was assessed at week 8 in patients with ≥2 lesions. The percentage change in each lesion was determined and the maximum difference calculated. A homogeneous response indicated that all the lesions for a patient had changed in the same direction with <30% difference between highest and lowest change. A low heterogeneous response indicated that all lesions changed in same direction, but that there was a ≥30% difference between the highest and lowest. For the homogeneous and low heterogeneous response categories, small changes (-10% to +10%) could be re-assigned to count as a change in the same direction. A high heterogeneous response indicated that at least one lesion underwent a ≤10% reduction *and* at least one other lesion underwent a ≥10% increase [6].

**Supplementary Figure 2: Radiological heterogeneity by sum of lesion diameters in patients with PR or SD at week 8.**  Percentage of patients with a sum of lesion diameters below (blue bars) or above the median size (orange bars) at week 8 which fall into the homogeneous, low heterogeneous or high heterogeneous radiological response categories. The difference between the two groups was not significant by Pearson’s Chi-Square test (*P*=0.17).