Optimizing the size variation threshold for the CT evaluation of response in metastatic renal cell carcinoma treated with sunitinib

R. Thiam1,2,3, L. S. Fournier1,2, L. Trinquart3, J. Medioni4, G. Chatellier2, D. Balvay2, B. Escudier5, C. Dromain6, C. A. Cuenod1,2* & S. Oudard4,7

1Department of Radiology, Georges Pompidou European Hospital, Assistance Publique Hôpitaux de Paris, Paris Descartes University; 2INSERM U970-laboratoire de Recherche en Imagerie team 2, Assistance Publique Hôpitaux de Paris, Paris Descartes University; 3Epidemiology and Clinical Research Unit, INSERM Centre d’Investigation Épidémiologique 4; 4Department of Oncology, Georges Pompidou European Hospital, Assistance Publique Hôpitaux de Paris, Paris Descartes University, Paris; Departments of 5Medical Oncology and 6Radiology, Gustave Roussy Institute, Villejuif and 7Biological Immunology Unit, Georges Pompidou European Hospital, Ecole Nationale vétérinaire d’Alfort EA 4054, Assistance Publique Hôpitaux de Paris, Paris Descartes University, Paris, France

Received 15 May 2009; revised 31 July 2009; accepted 24 August 2009

Background: In metastatic renal cell carcinoma (mRCC), antiangiogenic treatments rarely achieve a reduction of −30% in the sum of longest diameters (SLD) of target lesions required by RECIST for an ‘objective response’, although they objectively improve progression-free survival (PFS). We sought to determine a threshold for the computed tomography evaluation of these patients’ best reflecting patient outcome.

Patients and methods: In 334 mRCC patients treated with sunitinib, we tested thresholds from −45% to +10%. We classified patients as ‘responders’ when the best relative variation of the sum of longest diameters (D_{SLD}) reached the tested threshold and as ‘nonresponders’ otherwise. For each tested threshold, the median PFS of the two groups were compared. Receiver operating characteristic (ROC) analysis was also carried out among the 103 patients that progressed during follow-up. Finally, the ‘optimal’ threshold was retested on an independent cohort of 39 patients.

Results: The D_{SLD} threshold of −10% gave the most significant difference. It divided patients into 256 responders and 78 nonresponders (median PFS 11.1 and 5.6 months). The same −10% threshold was found using the ROC analysis. Results were confirmed on the external validation cohort.

Conclusion: A variation of −10% in the SLD accurately and rapidly identifies mRCC patients benefiting from sunitinib.

Key words: antiangiogenic, computed tomography, kidney cancer, metastatic renal cell carcinoma, progression-free survival, RECIST

background

The RECIST criteria [1, 2] are commonly used to evaluate the response of solid tumors to treatment in clinical trials. An ‘objective response’ is defined by a variation of at least −30% in the sum of the largest diameters of the target lesions. This −30% threshold, however, does not seem well adapted for the assessment of efficacy of antiangiogenic therapies. Indeed, tumor volume does not always decrease significantly according to RECIST criteria under antiangiogenic drugs as they act by inhibiting tumor vessels rather than by direct cytotoxic effect on tumor cells.

Antiangiogenic treatments have recently been tested in randomized phase III controlled trials involving metastatic renal cell carcinoma (mRCC) patients [3–6]. Sunitinib (Sutent®, Pfizer, New York, NY) [7] has been compared with interferon alpha demonstrating a longer median progression-free survival (PFS) in the antiangiogenic treatment arm (11 versus 5 months). This contrasts with the apparently low ‘response’ rate of 31% recorded according to RECIST 1.0, where a majority of patients (48%) were classified in the ‘stable disease’ group. Based on the improved PFS observed in this phase III study, sunitinib has been approved by USA and European authorities as a first-line treatment in mRCC [8].

In a clinical context, a patient with stable disease according to RECIST criteria, though he is not progressing, remains in an indeterminate category. Moreover, in cases where the threshold of −30% is reached, it often takes a long period of time to be achieved. A threshold that would accurately reflect patient outcome (biomarker) and was reached early on after beginning the treatment would avoid prolonged uncertainty for both clinicians and patients and would be helpful for a more efficient patient management.
Thus, the aim of this study was to identify, in a population of sunitinib-treated patients, a more efficient and earlier response criterion capable of separating patients for whom the treatment was likely to prolong PFS from those for whom it would not.

patients and methods

We conducted a retrospective sub-analysis of the international multicenter phase III randomized controlled trial [7] that compared the vascular endothelial growth factor receptor tyrosine kinase inhibitor, sunitinib, with first-line interferon alpha therapy in 750 patients with mRCC (375 patients in the sunitinib arm). Sunitinib was administered orally at a dose of 50 mg for 4 weeks, followed by a 2-week rest, with a dose reduction (to 37.5 or 25 mg) in cases of intolerance. Contrast-enhanced chest, abdomen and pelvis computed tomographies (CT) were carried out to measure target lesions before and every 6 weeks during treatment. For this sub-analysis, we used the measurement data obtained by the independent central review [7]. The present analysis focused on the 334 sunitinib-treated patients who had at least one target lesion (for determination of the sum of longest diameters (SLD)) and also for whom complete target lesion measurements were available. Eleven patients did not have a target lesion (only nontarget) and six patients had missing follow-up data concerning target lesion measurements within the database and were excluded from this study.

search for an optimal threshold

evaluation of response to treatment according to reference RECIST 1.0 criteria. According to the conventional RECIST 1.0 criteria, patients were separated into three groups, with the variation in the sum of the largest diameters (ASLD) classified below –30% (partial response), between –30% and +20% (stable disease) and above +20% (progressive disease). The best ASLD corresponded to the largest reduction in SLD observed during the course of the treatment as compared with the baseline SLD.

PFS was estimated using the Kaplan–Meier method. Progression and withdrawal or death were considered censoring events.

identification of the best threshold for ASLD by the Kaplan–Meier method. To find an optimal ASLD-based definition of response to treatment, we used the Kaplan–Meier method and tested a series of thresholds from –45% to +10%, in 5% increments. For each tested threshold, patients with an ASLD smaller than this threshold were classified as ‘nonresponders’ and those with an ASLD larger than the threshold were classified as ‘responders’. We compared the median PFS defined by the Kaplan–Meier curves of each of the two subgroups.

Two subgroups were considered distinct if the ratio of median PFS was significantly different from 1, i.e. if the 95% confidence interval (CI) associated with this ratio did not span 1.

threshold evaluation by receiver operating characteristic analysis. A different strategy involving receiver operating characteristic (ROC) curves was also used to determine the best threshold in the subgroup of 103 patients who had progressed under sunitinib during follow-up. These patients were classified into two groups according to whether they progressed before or after 5 months following inclusion in the phase III trial. This 5-month cut-off was chosen because it corresponded to the median PFS found in patients treated with interferon [7]. The ROC curve was used to identify the ASLD threshold with maximum sensitivity and specificity to differentiate patients benefiting from the treatment, with a prolonged PFS (>5 months), from patients who did not (<5 months).

The positive likelihood ratio (LR+) was calculated to analyze the concordance between the classifications according to the PFS at 5 months and to the best ASLD when using the new threshold versus the RECIST 1.0 criteria (~30%).

delay to reach response for the new threshold and the RECIST 1.0 threshold of ~30%

to identify the time of first occurrence of the threshold crossing, the numbers (percent) of patients reaching the new threshold or the conventional ~30% threshold were calculated for each cycle.

external confirmation on an independent cohort

The optimized threshold identified on the reference population (training set) was tested for confirmation on an independent cohort of 39 patients (age 56 ± 11 years) with mRCC issued from an open-label multicenter phase II trial [9]. The patients were on continuous sunitinib treatment at a dose of 37.5 mg/day. These patients were also followed by a contrast-enhanced chest, abdomen and pelvis CT to measure target lesions before and every 8 weeks during treatment. The same statistical analyses were carried out on this independent cohort as in the training set, comparing the median PFS of the two subgroups defined by the ‘optimized threshold’ and the RECIST 1.0 ~30% threshold and carrying out the ROC analysis in the 34 patients who progressed during follow-up.

Statistical analyses were done with SAS software v8.2 and GraphPad Prism v5.

results

search for an optimal threshold

evaluation of response to treatment according to reference RECIST 1.0 criteria. The best ASLD of target lesions defined using the RECIST 1.0 criteria was below ~30% in 146 (44%) of the 334 patients, between ~30% and +20% in 184 patients (55%) and above +20% in 4 patients (1%). At the cut-off date of the study, one hundred and three patients (31%) had progressed during follow-up.

Table 1 shows the median PFS and their ratio, in the two groups of patients defined using ASLD thresholds from ~45% to +10%. The thresholds of ~10%, ~5% and 0% all separated the population into two statistically distinct groups of responders and nonresponders. However, ~10% seemed the most clinically relevant threshold since there could be a risk that small changes in size <10% might be due only to measurement variability instead of a true change due to therapy.

With the ~10% threshold, patients whose best ASLD was superior to ~10% (n = 78) had a median PFS of 5.6 months (95% CI 3.0 to ∞), while those whose best ASLD was inferior to ~10% (n = 256) had a median PFS of 11.1 months (95% CI 10.3–12.1), thus deriving a clinical benefit from the treatment. The ratio of the two median PFS obtained with the ~10% threshold was 2.0 (95% CI 1.3–2.7) (Figure 1A). In contrast, the ratio of the two median PFS obtained with the RECIST 1.0 threshold of ~30% did not yield two significantly distinct groups (ratio 1.4, 95% CI 0.7–2.0) (Figure 1B).

threshold evaluation by ROC analysis. In the 103 patients in whom the effective PFS was available during follow-up, the ROC curve analysis yielded an optimal threshold of ~10% for the best ASLD (Figure 2), with a sensitivity of 52% and a specificity of 95%.

When crossing the classifications based on the PFS of 5 months and the threshold of ~10% or the RECIST 1.0 threshold in these patients (Table 2), the ~10% threshold had a negative predictive value of 94%. In contrast, patients classified...
as ‘stable’ according to RECIST 1.0 were truly indeterminate since they were as likely to have benefited from the treatment (PFS £ 5 months) as not (LR+ 1.4, 95% CI 1.0–2.0).

delay to reach response for the new threshold and the RECIST 1.0 threshold of –30%

In patients classified as responders according to the –10% threshold, the threshold was reached as soon as the first cycle of treatment in 73% of cases, whereas only 19% of patients responding according to the –30% threshold reached it during the first cycle (Table 3). The –10% and –30% thresholds were reached, respectively, in 93% and 64% of cases after the second cycle of treatment. Thus, the responder status was detected earlier with the threshold of –10% than with the conventional RECIST 1.0 threshold of –30%.

external confirmation on an independent cohort

The results in the external test set matched those observed in the training set.

The previously optimized –10% threshold divided the patients into two statistically distinct groups (Figure 3A), with median PFS of 2.6 and 8.1 months each (ratio 3.1, 95% CI 2.7–3.5), whereas the two groups obtained using the –30% threshold (Figure 3B) had median PFS of 7.9 and 6.5 months each, which were not statistically distinct (ratio 1.2, 95% CI 0.7–1.7).

In the patients who progressed under sunitinib during follow-up, the ROC curve analysis yielded the same optimal threshold of –10%, with a sensitivity and specificity of 50% and 100%, respectively.

When crossing the classifications based on the PFS of 5 months and the ASLD threshold of –10% or the reference RECIST 1.0 criteria, once again the RECIST 1.0 criteria of stable disease had poor ability to discriminate nonresponders (PFS £ 5 months) from responders (PFS > 5 months). Patients in the stable disease group were as likely to have benefited from sunitinib as not (LR+ 1.5, 95% CI 0.9–2.1).

Overall, 82% of patients under sunitinib treatment in the independent cohort could be considered as responders when

---

### Table 1. Median PFS for patients with the best variation of sum of longest diameters (ASLD) above and before the tested threshold and ratio of the median PFS, when testing various threshold values in the training set of 334 patients treated with sunitinib

<table>
<thead>
<tr>
<th>Threshold for the best variation of SLD (%)</th>
<th>Median PFS for patients above the threshold (months)</th>
<th>Median PFS for patients below the threshold (months)</th>
<th>Ratio of the median PFS (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>–45</td>
<td>9.6</td>
<td>12.1</td>
<td>1.3 (0.7–1.8)</td>
</tr>
<tr>
<td>–40</td>
<td>10.1</td>
<td>12.1</td>
<td>1.2 (0.6–1.8)</td>
</tr>
<tr>
<td>–35</td>
<td>8.5</td>
<td>11.5</td>
<td>1.3 (0.7–2.0)</td>
</tr>
<tr>
<td>–30</td>
<td>8.3</td>
<td>11.5</td>
<td>1.4 (0.7–2.0)</td>
</tr>
<tr>
<td>–25</td>
<td>7.5</td>
<td>11.5</td>
<td>1.5 (0.9–2.2)</td>
</tr>
<tr>
<td>–20</td>
<td>7.3</td>
<td>11.5</td>
<td>1.6 (0.9–2.3)</td>
</tr>
<tr>
<td>–15</td>
<td>7.3</td>
<td>11.1</td>
<td>1.5 (0.8–2.2)</td>
</tr>
<tr>
<td>–10</td>
<td>5.6</td>
<td>11.1</td>
<td>2.0 (1.3–2.7)</td>
</tr>
<tr>
<td>–5</td>
<td>4.1</td>
<td>11.0</td>
<td>2.7 (2.0–3.3)</td>
</tr>
<tr>
<td>0</td>
<td>1.5</td>
<td>11.0</td>
<td>7.3 (6.7–7.9)</td>
</tr>
<tr>
<td>+5</td>
<td>1.4</td>
<td>11.0</td>
<td>7.9 (7.3–8.4)</td>
</tr>
<tr>
<td>+10</td>
<td>1.4</td>
<td>11.0</td>
<td>7.9 (7.4–8.3)</td>
</tr>
</tbody>
</table>

PFS, progression-free survival; SLD, sum of longest diameters; CI, confidence interval; the threshold identified as optimal is in bold characters.
applying the \(-10\%\) threshold versus \(51\%\) when applying the RECIST 1.0 threshold of \(230\%\).

discussion

The RECIST 1.0 criteria used to assess the impact of chemotherapy on solid tumors is often criticized when used in the evaluation of response to targeted therapies such as antiangiogenic drugs. Indeed, the RECIST 1.0 criteria were developed as a statistical tool to evaluate response in large populations participating in cytotoxic drug trials [2]. The new version RECIST 1.1 [10] added modifications to the previous version RECIST 1.0 for lymph node evaluation [11] and for progression but maintained a \(230\%\) threshold-defining response. The low rates of tumor shrinkage are a problem specific to targeted therapies such as antiangiogenic drugs, and we believe that the detection of tumor response with these drugs must be addressed specifically. Techniques such as functional imaging seem promising in this context [12–16]; however, they are not yet ready to be clinically implemented.

Therefore, alternative radiological classification tools seem necessary to correctly evaluate the efficacy of antiangiogenic agents in patients.

Our study shows that in a population of 334 mRCC patients treated with the antiangiogenic drug sunitinib in an international multicenter phase III trial [7], a threshold of \(-10\%\) for the relative change in the sum of the largest tumor diameters of

Table 2. Concordance between the classifications based on the progression-free survival of 5 months and the classifications based on the variation of SLD when using the threshold at \(-10\%\) or \(-30\%\), in the 103 patients under sunitinib who progressed during follow-up.

Table 3. Number of patients reaching threshold per treatment cycle in the 334 patients treated with sunitinib, when using the \(-10\%\) and RECIST 1.0 \(-30\%\) thresholds.

SLD, sum of longest diameters; LR+, positive likelihood ratio; CI, confidence interval.

Three hundred and thirty-four patients were included at baseline, but the number of patients decreased at each cycle, either because they did not have their examination on that cycle or because they were censored.

Figure 2. Receiver operating characteristic curves to identify the best threshold in the subgroup of 103 patients who had progressed under sunitinib during the follow-up. Patients were classified into two groups according to whether they progressed before or after 5 months.

Table 2.

<table>
<thead>
<tr>
<th>Training set</th>
<th>Progression-free survival ≤5 months</th>
<th>&gt;5 months</th>
<th>Total</th>
<th>LR+ (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threshold at (-10%) or more:</td>
<td>33</td>
<td>2</td>
<td>35</td>
<td>10.5 (3.1–38.5)</td>
</tr>
<tr>
<td>‘nonresponders’</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than (-10%):</td>
<td>30</td>
<td>38</td>
<td>68</td>
<td>0.5 (0.4–0.6)</td>
</tr>
<tr>
<td>‘responders’</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>63</td>
<td>40</td>
<td>103</td>
<td></td>
</tr>
<tr>
<td>Threshold at (-30%) or more:</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>‘progressors’</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(-30%) to (+20%):</td>
<td>48</td>
<td>22</td>
<td>70</td>
<td>1.4 (1.0–2.0)</td>
</tr>
<tr>
<td>‘stables’</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(-30%) or less:</td>
<td>11</td>
<td>18</td>
<td>29</td>
<td>0.4 (0.2–0.7)</td>
</tr>
<tr>
<td>‘responders’</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>63</td>
<td>40</td>
<td>103</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.

<table>
<thead>
<tr>
<th>Training set</th>
<th>Cycle 1, n (%)</th>
<th>Cycle 2, n (%)</th>
<th>Cycle 3, n (%)</th>
<th>Cycle 4, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threshold for the best variation of sum of longest diameters</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than (-10%)</td>
<td>186 (73)</td>
<td>237 (93)</td>
<td>251 (98)</td>
<td>256 (100)</td>
</tr>
<tr>
<td>(-30%) or less</td>
<td>28 (19)</td>
<td>93 (64)</td>
<td>125 (86)</td>
<td>140 (96)</td>
</tr>
<tr>
<td>(-30%) to (+20%):</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘nonresponders’</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘stables’</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘progressors’</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>214</td>
<td>330</td>
<td>376</td>
<td>400</td>
</tr>
</tbody>
</table>
target lesions is optimal to separate the patients into two groups according to the benefit they derived from the treatment (median PFS of 11.1 versus 5.6 months). Moreover, responders defined by this new threshold had a PFS statistically equivalent to that of patients with partial response according to the RECIST 1.0 reference, confirming that these patients constitute a statistically coherent population with a similar clinical outcome under treatment by sunitinib. In this analysis, P value was not used to separate the two subgroups because we were interested in the median duration and not survival probability.

Responders were identified earlier during the course of the treatment with the −10% threshold (reached by 73% of patients after only 6 weeks of treatment) than with the −30% RECIST 1.0 threshold (reached by 64% of patients after 12 weeks of treatment). Furthermore, the −10% threshold had a very high specificity, correctly detecting nonresponders who might benefit from a change of treatment and avoid unnecessary toxic effects.

It should be noted that our study population is large and homogeneous ensuring robust size analysis and that all the radiological data were reviewed by an independent panel. Two independent statistical methods—Kaplan–Meier survival analysis and ROC curves—yielded the same −10% threshold for the best variation of the sum of the largest tumor diameters. Moreover, these results were confirmed on an external population of mRCC patients from a phase II trial using continuous therapy with sunitinib [9]. In this independent set of patients, the −10% threshold yielded a high rate of response (82%), and this threshold defined two distinct populations of patients, with different outcomes. It is interesting to note that the −10% threshold proved to be appropriate to evaluate response to sunitinib in this new population, even though they received the drug at a different regimen than patients in the reference cohort.

Our study is the first, to our knowledge, to evaluate threshold optimization for treatment response in mRCC patients under antiangiogenic drugs. A previous study in patients with gastrointestinal stromal tumors treated with imatinib, a novel targeted agent [17], showed that a 10% reduction in size or a 15% reduction in tumor intensity on contrast-enhanced CT was predictive of slower progression and was sufficient to consider that the patient was benefiting from the treatment.

The −10% threshold has the advantage of categorizing patients into only two groups, whereas RECIST 1.0 categorizes patients into three groups (partial response, stable disease and progressive disease). The stable disease category of RECIST 1.0 pools without discrimination patients who will have a PFS prolonged by therapy and patients who will not derive any clinical benefit from it, thus providing no information on treatment efficacy. In our study, we showed that patients in the stable category had as much chance of having a prolonged PFS >5 months as not. Indeed, the RECIST 1.0 classification was developed for evaluation of large clinical trials, and not for individual day-to-day decision making in patients, on whether to pursue or interrupt a drug or change the treatment dose [18, 19].

Based on previous studies, we chose to define patient benefit as having a PFS >5 months because this value represents the median PFS of patients under interferon, the reference treatment used in the phase III trial of which our study was a sub-analysis [7]. This was indeed an arbitrary decision which seems reasonable in regard to the 2.8- and 5.5-month PFS of patients, respectively, under placebo or sorafenib [3]. Moreover, secondary tests with other PFS thresholds between 3 and 7 months gave similar results.

However, the use of thresholds for changes in tumor size to assess the treatment response of solid tumors has its own limits since two patients with very close responses, for example −9% and −11%, will be categorized into two separate groups (nonresponder versus responder). However, this is already the case with any categorizing strategy including RECIST 1.0 used in clinical trials since patients immediately above or under the +20% threshold for progression will also be categorized as two distinct responses. In a clinical context, however, oncologists use thresholds to guide their therapeutic decisions but take also into account other parameters, such as clinical evaluation of the patient and biology.

Intraobserver and interobserver reproducibility is an issue in tumor size assessment [20–22]. The reproducibility of
CT-based tumor size measurement was assessed by Hopper et al. [23] showing an interobserver variability of 3%–15%, but in this study the measures were done manually, not with an electronic caliper found on workstations, which is the current method. Among the thresholds that we tested, the 0% (ratio 7.3; 95% CI 6.3–8.2) and −5% values (ratio 2.7; 95% CI 2.0–3.3) also separated the population into two distinct groups when using the Kaplan–Meier method, but there was a risk that they would fall within the margin of interobserver variability. We therefore more conservatively prefer a threshold of −10%, to take into account the variability induced by repeated measures, slice positioning, movement artifacts, the choice of target lesions and interobserver and intraobserver variability. It should also be noted that the ROC analysis also yielded −10% as the best threshold. However, we realize that in cases of patients with only one target lesion, a 10% change may turn out to be difficult to measure. Indeed, for a lesion measuring 10 mm, a 10% change represents only 1 mm. It is unlikely that a method of measurement could have an accuracy of 1 mm; therefore, one solution to avoid making conclusions on such small values would be to adopt the same recommendations as RECIST 1.1 (though concerning progression in their case), which are that for a change to be significant it must have an absolute value >5 mm. It may also be important to try to select several target lesions whenever possible. However, as stated above, in a clinical trial, a threshold is just a statistical tool to demonstrate treatment efficacy, and in a clinical context, the threshold should only serve as a guide for the clinician rather than an absolute truth dictating his decisions.

Finally, in this study, we did not repeat the same analysis for the threshold using the new RECIST 1.1 criteria [10, 11] on this retrospective study initially analyzed following RECIST 1.0 strategy. A new reading of the original images or further studies would have to be carried out to determine whether this threshold remained a good indicator of patient outcome when using the new RECIST 1.1 guidelines.

Conclusions

A relative reduction of −10% in the sum of the largest tumor diameters appears to be a reliable threshold for identifying patients with mRCC who are benefiting from antiangiogenic treatment with sunitinib and for aiding the oncologists’ decisions. Patients reach this threshold earlier than the RECIST 1.0 −30% threshold, allowing a quicker evaluation of treatment efficacy. Further studies will be necessary to confirm these findings and to evaluate its potential use for clinical trials or for individual patient follow-up. A similar methodology could be used for optimizing the threshold for other type of cancers and antiangiogenic treatments.

Acknowledgements

We thank Isan Chen and Rémy Defrance (Pfizer Inc.) for providing us with the database; Guy Frija for his help in the conduct of the study and Joelle Bauvillard, Catherine Boulongne and Catherine Cornuault for their help with the data collection.

References