Eur Radiol (2016) 26:2798–2807
DOI 10.1007/s00330-015-4082-8

MOLECULAR IMAGING

The radiogenomic risk score stratifies outcomes in a renal cell cancer phase 2 clinical trial

Neema Jamshidi[1] · Eric Jonasch[2] · Matthew Zapala[1,3] · Ronald L. Korn[4] · James D. Brooks[5] · Borje Ljungberg[6] · Michael D. Kuo[1]

Received: 18 July 2015 / Revised: 30 September 2015 / Accepted: 23 October 2015 / Published online: 11 November 2015

Abstract
Objectives To characterize a radiogenomic risk score (RRS), a previously defined biomarker, and to evaluate its potential for stratifying radiological progression-free survival (rPFS) in patients with metastatic renal cell carcinoma (mRCC) undergoing pre-surgical treatment with bevacizumab.
Methodology In this IRB-approved study, prospective imaging analysis of the RRS was performed on phase II clinical trial data of mRCC patients (n=41) evaluating whether patient stratification according to the RRS resulted in groups more or less likely to have a rPFS to pre-surgical bevacizumab prior to cytoreductive nephrectomy. Survival times of RRS subgroups were analyzed using Kaplan-Meier survival analysis.
Results The RRS is enriched in diverse molecular processes including drug response, stress response, protein kinase regulation, and signal transduction pathways ($P<0.05$). The RRS successfully stratified rPFS to bevacizumab based on pre-treatment computed tomography imaging with a median progression-free survival of 6 versus >25 months ($P=0.005$) and overall survival of 25 versus >37 months in the high and low RRS groups ($P=0.03$), respectively. Conventional prognostic predictors including the Motzer and Heng criteria were not predictive in this cohort ($P>0.05$).
Conclusions The RRS stratifies rPFS to bevacizumab in patients from a phase II clinical trial with mRCC undergoing cytoreductive nephrectomy and pre-surgical bevacizumab.
Key Points
- *The RRS SOMA stratifies patient outcomes in a phase II clinical trial.*
- *RRS stratifies subjects into prognostic groups in a discrete or continuous fashion.*
- *RRS is biologically enriched in diverse processes including drug response programs.*

Keywords Radiogenomics · Imaging surrogate · Renal cell carcinoma · Imaging biomarker · Bevacizumab

Electronic supplementary material The online version of this article (doi:10.1007/s00330-015-4082-8) contains supplementary material, which is available to authorized users.

✉ Michael D. Kuo
michaelkuo@mednet.ucla.edu

[1] Department of Radiological Sciences, University of California-Los Angeles, David Geffen School of Medicine, Los Angeles, CA, USA

[2] Department of Genitourinary Medical Oncology, The University of Texas M.D. Anderson Cancer Center, Houston, TX, USA

[3] Department of Radiology, University of California-San Diego, San Diego, CA, USA

[4] Scottsdale Medical Imaging, Scottsdale, AZ, USA

[5] Department of Urology, Stanford University School of Medicine, Stanford, CA, USA

[6] Department of Surgical and Perioperative Sciences, Urology and Andrology, Umea Hospital, Umea, Sweden

Abbreviations

RRS	radiogenomic risk score
SOMA	surrogate of molecular assay
WHO	World Health Organization
SPC	supervised principal component
ccRCC	clear cell renal cell carcinoma
mRCC	metastatic renal cell carcinoma
GO	gene ontology
MSKCC	Memorial Sloan Kettering Cancer Center
RECIST	response evaluation criteria in solid tumours

Introduction

Clear cell renal cell carcinomas (ccRCCs) are the most common and lethal form of renal malignancies [1]. Once they become metastatic, treatment becomes difficult and mortality sharply increases, even in the face of new anti-angiogenesis therapies, with a median survival of only 1 to 2 years [1]. Thus, identification of patients likely to obtain durable benefit from anti-angiogenesis therapy once they become metastatic remains a goal of paramount importance.

Radiogenomics, the multi-scale integration of biological data with multi-parametric imaging data, has been developed with the goal of developing new tools to provide insight into the relationships between imaging, cellular and subcellular data [2, 3]. To date, it has been used to identify robust non-invasive biomarkers in multiple cancer types using different imaging modalities for various uses, including predicting complex patterns of gene expression, prognosis, and treatment response [4–11]. Radiogenomics thus affords an opportunity to identify patients based on their genomic and physiological alterations in a noninvasive and serial manner, and, thus, is potentially of paramount clinical relevance to the design and implementation of clinical trials. We recently showed that quantitative tissue-based molecular assays in clear cell renal carcinoma can be efficiently translated into semi-quantitative image based, non-invasive surrogates of molecular assays (SOMAs) using clinical image data from contrast-enhanced computed tomography (CT) scans [15].

There are numerous instances in medicine where therapies or diagnostics designed around one specific purpose, upon further testing and evaluation, have been shown to be successful in broader applications within the same disease or even in adjacent domains [12–14]. However, the broader question of whether multi-feature radiogenomic biomarkers can be similarly designed, and then successfully tested and scaled beyond their initial prescribed indications remains unanswered. Specifically, it is unknown whether the biological information encapsulated by a radiogenomic SOMA is sufficient to capture other clinically relevant parameters beyond its initial intended purpose.

We hypothesized that quantitative, multi-featured, non-invasive, image-based molecular assays (SOMA) could be phenotypically characterized and, based on this characterization, further extended to predict clinical phenotypes beyond their original indication. Herein, we characterize and define properties of the newly discovered RRS and then leverage this information to test its ability to stratify patients based on clinical outcomes in a phase II clinical trial of pre-surgical bevacizumab treatment in metastatic RCC patients.

Materials and methods

Patients and materials

Patients from the training/validation set

Fresh frozen tissue, clinical outcomes, and matched pre-operative diagnostic contrast-enhanced CT scans were acquired from Umea University Hospital in accordance with local institutional Umea University Hospital institutional review boards and in compliance with the Declaration of Helsinki. The training set (n=70) was composed of patients treated between 1994 and 2003. The validation set (n=77) was composed of patients treated between 2000 and 2007; of the 77 patients in the validation set, 70 had sufficient tissue for gene expression analysis. Gene expression profiling, SPC risk score, and RRS generation was performed on a total 140 patients; descriptions of these two cohorts are fully detailed in [15].

Phase II clinical trial patients

Patients were enrolled in the single centre phase II clinical trial (NCT00113217) at the MD Anderson Cancer Center between 2005 and 2008 in order to assess clinical outcomes of patients with newly diagnosed, untreated mRCC who received the anti-angiogenesis therapy bevacizumab (Avastin, Genentech, South San Francisco, CA) before undergoing cytoreductive nephrectomy. Fifty patients with histologically confirmed mRCC were ultimately analyzed in the initial phase II trial; however, of these, 42/50 (84 %) completed pre-surgical bevacizumab and cytoreductive nephrectomy per study protocol. One additional patient was removed due to incomplete imaging data resulting in a total of 41 patients analyzed in this study. None of the patients in the phase II clinical trial were in the original training and validation data sets [15]. Clinical endpoints from the trial used in this analysis included time to disease progression, and overall survival per study protocol; all tumour measurements were made using RECIST 1.0. Complete details of the trial and study cohort are provided in [16]. The Memorial Sloan Kettering Cancer Center (MSKCC, also known as the Motzer criteria) and Heng criteria, a general RCC prognostic model and an advanced-stage RCC specific anti-angiogenesis predictive model, respectively [17], were calculated for each patient as previously described [18–20].

CT imaging, RRS image analysis and interobserver variability

All CT scans were performed with renal mass protocols using a multi-detector CT scanner (HD750, GE Healthcare) according to the study protocol. Initially, an unenhanced CT scan of the liver and kidneys (top of the liver to below kidneys) was performed followed by a

100-mL bolus of non-ionic contrast material (Omnipaque-350, GE Healthcare) administered via an automatic injector at a rate of 5 ml/s intravenously with a delay of 40 seconds followed by 30 ml of normal saline at 5 ml/sec. A second injection of 50 ml of non-ionic contrast material (Omnipaque-350, GE Healthcare) was administered via an automatic injector at a rate of 3 ml/s intravenously at a 10-minute delay followed by 30 ml of normal saline at 3 ml/sec. The images were obtained at 5 mm and reconstructed at 2.5 mm and 1.25 mm. Axial and multi-planar reformatted images in sagittal and coronal planes were generated for interpretation.

All patients' registration CT images were evaluated on a DICOM viewer workstation (OsiriX 64-bit) independently and resolved in consensus across the 4 RRS features by two board certified radiologists (MK, RK, 6 and 12 years, respectively).

Radiogenomic characterization of the radiogenomic risk score

Phenotype landscape map analysis

The four RRS image features: pattern of tumour necrosis (trait 17), tumour transition zone (trait 26), tumour-parenchyma interaction (trait 30), and tumour-parenchyma interface (trait 31) imaging features were interpreted and the RRS classification was applied using the equation, $Y=\beta_1X_1+\beta_2X_2+\beta_3X_3+\beta_4X_4+\epsilon$, where ϵ=-0.187, β_1=0.03591, X_1=pattern of tumour necrosis (PTN), β_2=0.113, X_2=tumour transition zone (TTZ), β_3=-0.124, X_3=tumour-parenchyma interaction (TPIa), β_4=-0.08016, X_4=tumour-parenchyma interface (TPIf) as recently described [15]. A colour-coded heat map was constructed visualizing the relationships between the range of different individual trait score values, their corresponding composite RRS values and the binary RRS score (high/low RRS). Statistical properties and characteristics of the RRS were calculated from the coordinate map.

Biological and genomic characterization

The set of genes that were significantly associated with the RRS and shared between the training and validation sets (n=140) [15] were identified through calculation of the Spearman correlation coefficient with the null distribution determined from 5000 permutations of the calculated correlation coefficients without replacement and significance criteria set at $P<0.05$. We then assessed for enrichment of gene ontology (GO) categories including *Biological Process*, *Cellular Component*, and *Molecular Function* ontologies using GoMiner [21] ($P<0.05$).

Statistical analysis

To define the continuous relation between the RRS and the 10-year risk of ccRCC disease-specific survival, the data were fitted by a time-varying, piecewise, log-hazard ratio model with the RRS score and its quadratic term included as covariates. The 10-year rate of disease-specific survival was then estimated by a Breslow-type function [22]. Kaplan-Meier plots were generated for survival outcomes analyses with log-rank tests. Interobserver variability was assessed and reported using Cohen's kappa statistic [23]. Multivariate Cox regression was performed to assess significance and dependence of the RRS, MSKCC criteria and Heng criteria.

Analyses were performed using R (http://www.r-project.org/), Matlab (The MathWorks, Natick, MA), SPSS (IBM, Armonk, NY), Statistica (StatSoft, Tulsa, OK), and GoMiner (http://discover.nci.nih.gov/gominer).

Results

The RRS is a linear combination of four stage- and grade-independent imaging traits: pattern of tumour necrosis (PTN), TTZ, TPIa, and tumour-parenchyma interface (TPIf) with a binary cutoff value into high and low risk groups of -0.01318. The individual traits that comprise the RRS are linearly independent as measured by Pearson correlations (Table E1) and are independent of WHO grade and stage, as previously described [15].

The phenotypic landscape map of the RSS radiophenotype

In order to better understand the relationship between the different imaging features, how they affect one another, and how disease-specific survival is affected by these interactions, a phenotypic landscape of the RRS was constructed (Fig. 1).

The RRS radiophenotypic landscape can be viewed as a coordinate map of the predicted clinical outcome overlain on the phenotypes, and reveals how a patient's risk is affected when shifting from one location in the map (corresponding to a set of imaging features and an associated risk score) to another point in the image feature space. A total of seven transition zones, where a change in score of a single trait will cause the risk phenotype to switch from low to high risk or vice-versa, are observed (delineated by the *black line* in Fig. 1). Conversely, the landscape map also reveals areas of the map where changes in one, or even all, of the imaging trait score values will not alter the high- or low-risk group RRS designation (Figure E1). Thus, the radiophenotype landscape allows one to identify

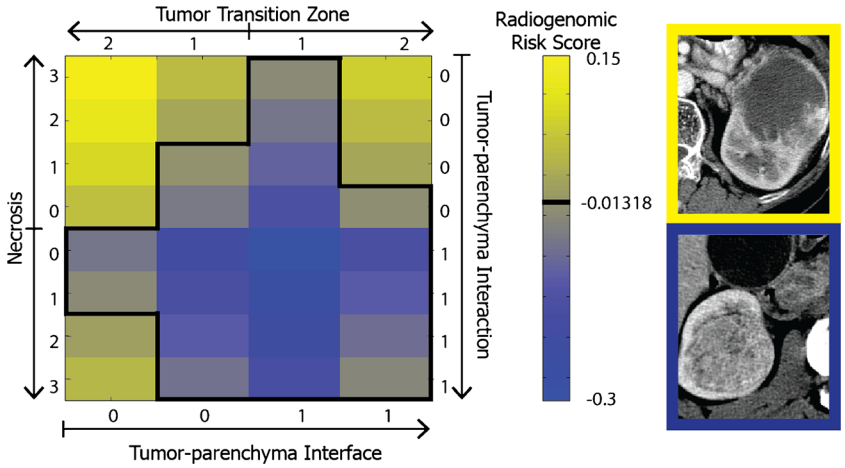

Fig. 1 Characterization of the radiophenotypic landscape of the RRS. The plot displays the range of all possible expression values for each of the four image traits (PTN, TTZ, TPIa, TPIf) that, in aggregate, define the composite risk phenotype; the *black boundary line* (risk score of -0.01318) highlights the transition points between the high- (*yellow*) and low-risk score (*blue*) groups. Exemplar images of the high- and low-risk score phenotypes are shown to the right (*yellow outline* = high-risk score and *blue outline* = low-risk score)

particular combinations of traits that are more or less likely to shift the predicted RRS risk scores.

The map reveals that TTZ and TPIa have the strongest impact on overall risk score consistent with their respective coefficients in the RRS equation. However, we observe that each of the four selected traits is independently relevant and can cause a shift in the predicted risk of the patient. If one assumes that a transition can occur between any two points in the map, then 66 % (21/32) of the transitions will lead to low-risk phenotypes and 34 % (11/32) transitions will lead to corresponding high-risk phenotypes. If the initial state is a low-risk phenotype, the odds of transitioning from a lower risk score to a high risk score is 0.52 (34/66). If however, the transitions of the RRS are sequential (i.e. given any state in the map in Fig. 1, the RRS changes by moving a single step horizontally or vertically), then the chance of transitioning from a low-risk phenotype to a high-risk phenotype or vice-versa is only 25 % (13/52).

The RRS as a continuous predictor of outcome

As a multivariate linear regression equation, the RRS is not restricted to use as a categorical classifier (high and low risk), but can also be used as a continuous predictor of disease-specific survival (Fig. 2). The likelihood of survival decreased continuously as the risk score increased. Two-sided confidence intervals for the likelihood of disease-specific death for the training and validation sets for both groups (low- and high-RRS groups) are±6.6 %, and are generally±3.2 % for RRSs of less than or equal to -0.01318 (low-risk score group) and±11.4 for RRS greater than or equal to -0.01318 (high risk score group).

Affect of the RRS on disease-specific survival within grade, stage, and WHO performance subgroups

Subgroup analysis of the RRS by grade, WHO performance status and stage reveals significant differences in disease-specific survival based on stratification of patients into high- and low-risk RRS score with high-RRS groups revealing significantly worse outcomes than the low-RRS groups: high-grade (grades 3/4) and low-grade (grades 1/2) patients (P= 0.006 and P<0.0001, respectively), WHO performance status (WHO performance 0/1, P<0.0001, WHO performance 2/3, P<0.0001). The RRS was also significant based on advanced stage (stage 3/4, P=0.007). Results are summarized in Supplementary Figure E2.

GO Analysis of the RRS

Collectively, 137 out of a total 30,057 uniquely measured genes (0.005 %) from the annotated probes were significantly associated with the RRS (P<0.05 by re-sampling of the Spearman correlation coefficient). GO enrichment used to characterize the pathway enrichment of molecular biological processes on these 137 RRS-associated transcripts revealed significant enrichment in 125 diverse cellular processes, including drug response, stress response, immune response, signalling phosphorylation pathways, protein kinase regulation, and hematopoietic and myeloid pathways (Table 1 and Table E2).

RRS stratifies outcomes in a phase II clinical trial of pre-surgical bevacizumab in mRCC

Given the evidence from this radiogenomic characterization of the RRS, which revealed an ability to stratify outcomes in

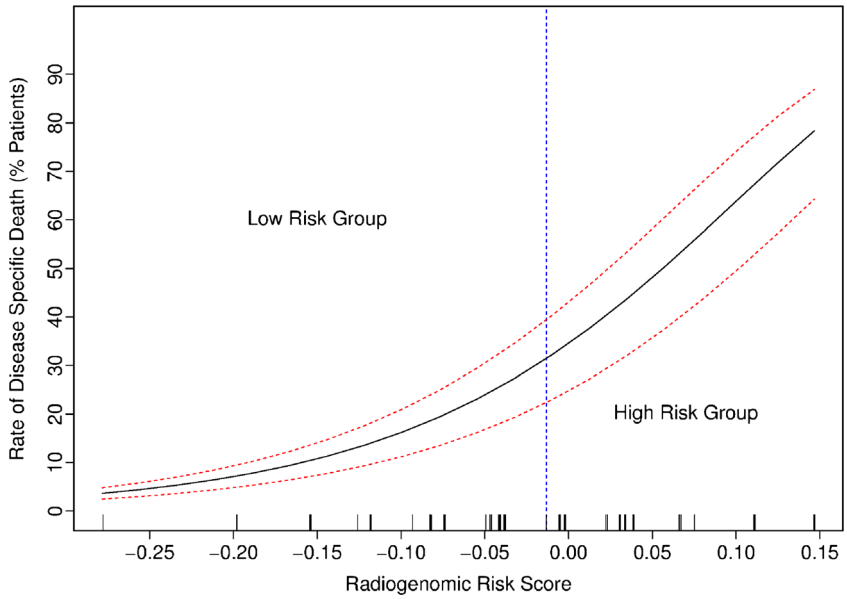

Fig. 2 Rate of disease-specific-death as a continuous function of the RRS. The continuous function was generated using a piecewise log hazard ratio model. The *dashed curves* indicate the 95 percent confidence interval. The rug plot above the x-axis shows the risk score for individual patients (n=140)

patients with advanced/metastatic disease (the target population of RCC anti-angiogenesis therapies; Figure E2), and similarly showed significant enrichment in protein kinase and drug response pathways (Tables 1 and E2), we prospectively evaluated the hypothesis that the RRS could stratify patients based on the radiologic progression-free survival (rPFS) to anti-angiogenesis therapy in mRCC patients in independent data from a phase II clinical trial of mRCC patients treated with the anti-angiogenesis drug bevacizumab prior to nephrectomy (see Jonasch et al [16] for full study details). We find that the RRS successfully stratifies patients based on both rPFS and overall survival in this cohort based on the pre-treatment CT appearance of the primary tumour (n=41) prior to initiation of anti-angiogenesis therapy (Fig. 3, *panels A* and *B*) with a median progression-free survival of 6 months compared to >25 months (P=0.005) in the high- and low-RRS groups, respectively, and overall survival of 25 months compared to >37 months (P=0.03) in the high- and low-RRS groups, respectively (Fig. 3, *panels A* and *B*). Representative images from patients with low- and high-RRS groups are highlighted in Fig. 3, *panels C* and *D*. There was good inter-observer agreement of the RRS as measured by Cohen's kappa statistic; 0.796±0.097 (4 /41 discrepancies resolved in consensus).

To our knowledge, there is currently no means of stratifying risk based on disease-specific survival across all stages of ccRCC that is also able to stratify rPFS to treatment with bevacizumab, a drug currently indicated for late-stage

Table 1 GO enrichment of the RRS. Subset of the GO enrichment terms based upon the set of genes that were significantly correlated with the RRS from the training and validation sets (P<0.05) that represent a "response" category to a biological process. The full set of 125 GO terms that were statistically significantly enriched is tabulated in Table E2

GO ID	Total	Change	*P*-value	Term
2526	4	2	0.0257	Acute inflammatory response
2263	3	2	0.0134	Cell activation involved in immune response
2366	3	2	0.0134	Leukocyte activation involved in immune response
2275	3	2	0.0134	Myeloid cell activation involved in immune response
48585	11	3	0.0333	Negative regulation of response to stimulus
80135	5	2	0.041	Regulation of cellular response to stress
48583	34	6	0.0183	Regulation of response to stimulus
9617	10	3	0.0253	Response to bacterium
42221	76	9	0.0417	Response to chemical stimulus
42493	16	4	0.0177	Response to drug
10033	44	7	0.0175	Response to organic substance
50896	118	13	0.0093	Response to stimulus
6950	75	12	0.0003	Response to stress
9611	41	6	0.0455	Response to wounding

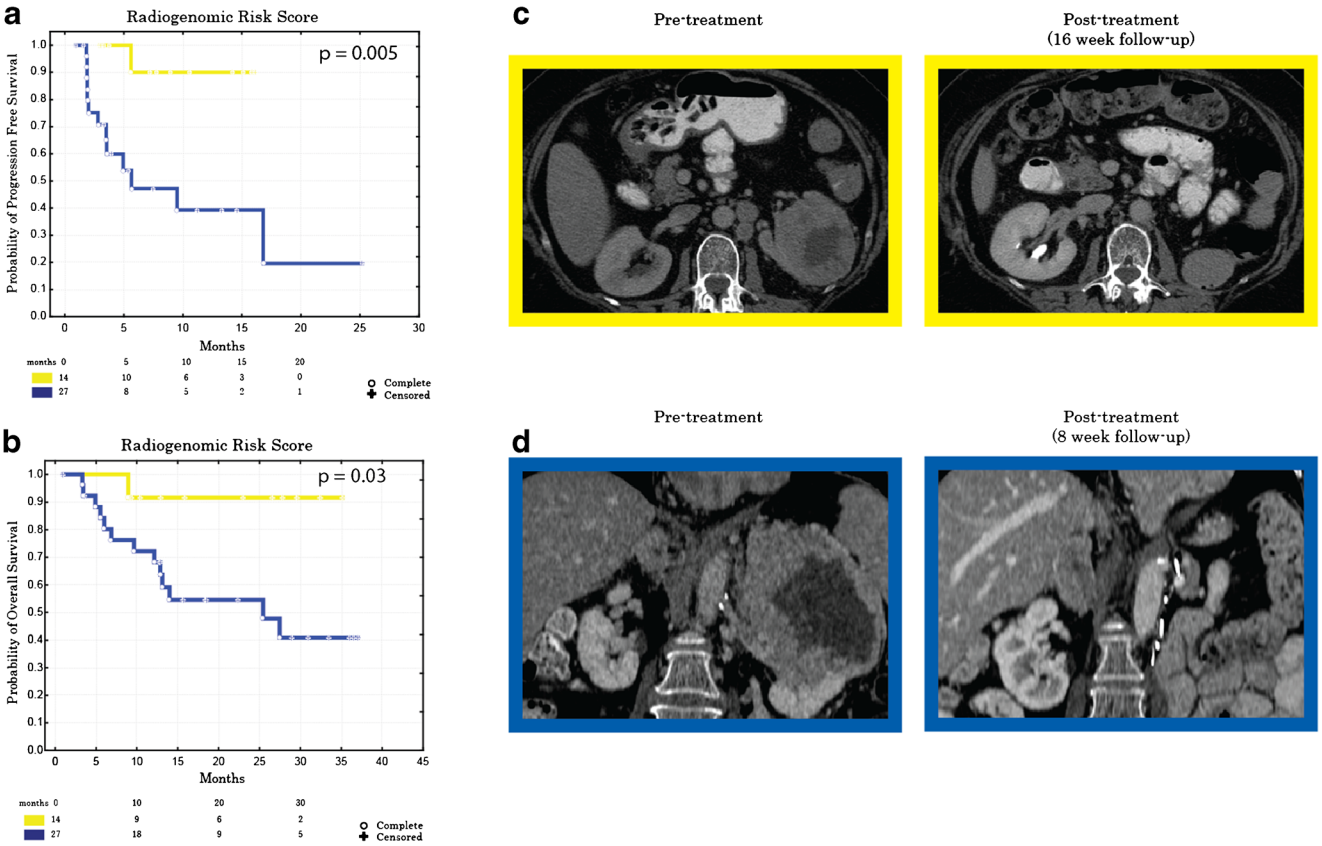

Fig. 3 The RRS predicts response to anti-angiogenesis therapy in a phase II clinical trial (n=41; A–F) [16]. Kaplan-Meier treatment response curves for high- and low-RRS groups in patients treated with bevacizumab are shown for progression-free survival (*A*) and overall survival (*B*), both of which are statistically significant at $P<0.05$. Pre-treatment and 16-week follow up images following nephrectomy and cytoreduction with bevacizumab for a 71-year-old female (RRS=-0.082) with retroperitoneal lymphadenopathy without evidence of progression at 231 days (*C*). Pre-treatment and 8-week follow up images following nephrectomy and bevacizumab with a metastatic right adrenal mass initially measuring 3.1 x 2.1 cm with evidence of disease progression at 8 weeks with increased size of the adrenal lesion to 5.2 x 2.6 cm for a 46-year-old male (RRS=0.066; *D*)

patients. However, we considered the possibility that other outcome predictors for late-stage RCC might also be able to demonstrate similar capabilities in this analysis. According to MSKCC [24] criteria, a general indicator of prognosis across all stages of RCC, progression-free survival was 17 months versus 6 months and greater than 37 months versus 13 months for overall survival ($P=0.10$ and $P=0.12$, respectively; Fig. 4, *panels A* and *B*). Assessment and stratification based on the Heng [19] criteria, an indicator of prognosis in patients treated with anti-angiogenesis therapies, resulted in progression-free survival of 17 months versus 9 months and greater than 37 months for overall survival (both groups; $P=0.33$ and $P=0.63$, respectively; Fig. 4, *panels C* and *D*). The lack of any statistically significant prognostic stratification in this study population suggests that the RRS tracks a phenotype independently of these clinical prognostic criteria (Fig. 4, *panels A–D*).

A composite portrait of high- versus low-RRS group patients is presented in Table 2, demonstrating demographic, clinical outcomes, objective response rates, and MSKCC and Heng criteria data in the phase II clinical trial dataset. Further, there were no differences between the high- and low-risk RRS group to suggest bias with respect to metastatic site involvement, primary tumour size or disease burden (Table 2). As the RRS, and MSKCC and Heng criteria are all prognostic indicators capable of providing a read-out from a single time point prior to initiation of treatment (as opposed to RECIST which measures tumour responses as captured by changes in tumour size over the arc of treatment) we evaluated a multivariate Cox model (Table 3) which showed that only the RRS was significant (hazard ratio [HR]=8.58, $P=0.04$), confirming that the RRS is an independent predictor of outcome in patients treated with bevacizumab in this cohort.

Discussion

There has been a push to identify biomarkers able to better stratify patients based on likely rPFS to therapy in the setting of early-phase clinical trials. Although clinical trials are increasingly incorporating tissue-based biomarkers [25–27], it is well-accepted that the cellular and genomic landscape of

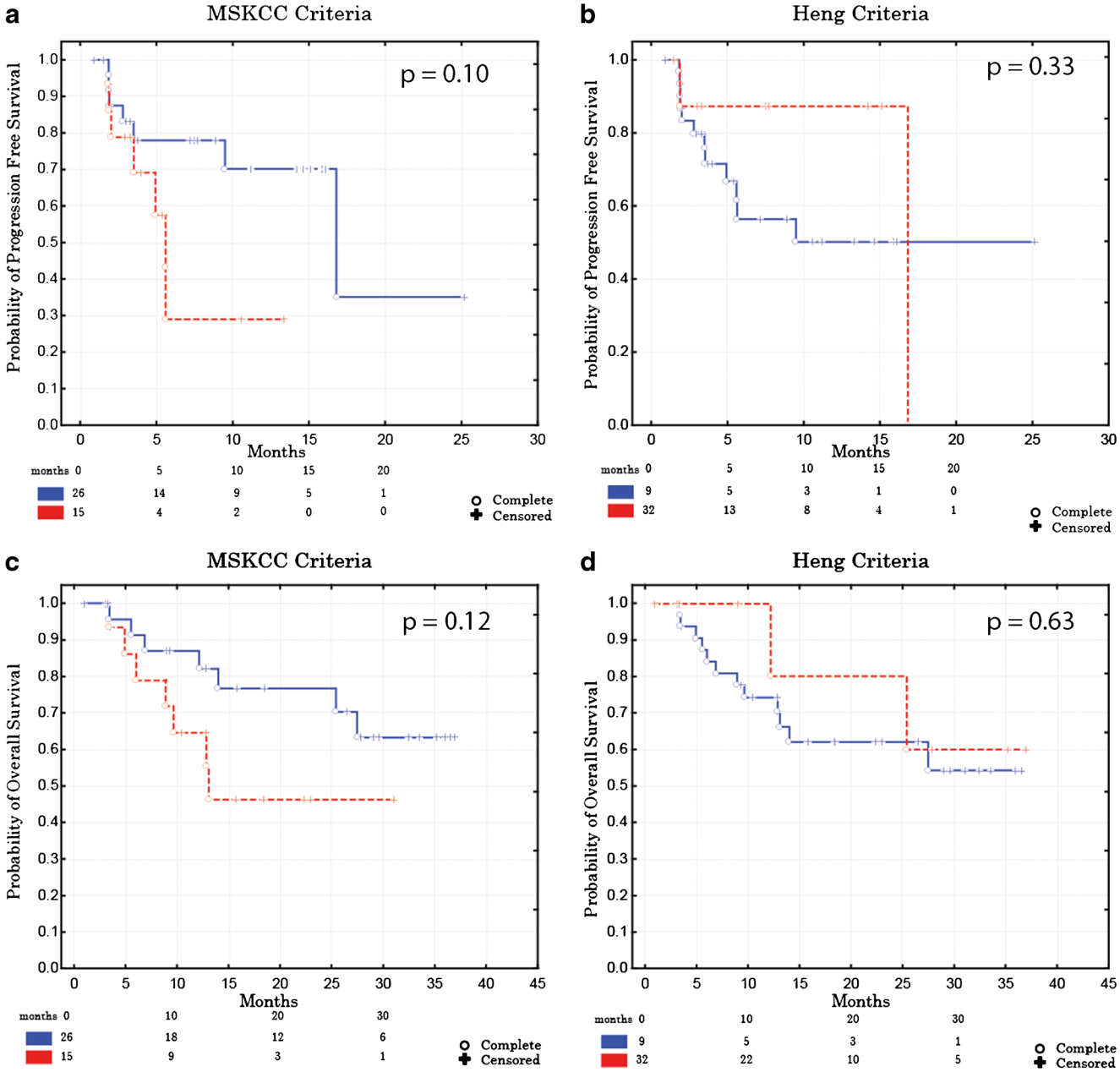

Fig. 4 Other prognostic criteria are not predictive to anti-angiogenesis therapy. Evaluation of MSKCC prognostic criteria in a phase II clinical trial cohort for progression-free survival (*A*), OS (*B*), and the Heng treatment response criteria for progression-free survival (*C*) as well as for overall survival (*D*; $P>0.05$)

tumours are highly heterogeneous, and, thus, subject to sampling bias, bringing to question their relevance [4, 28, 29]. Herein we show that a rationally constructed radiogenomic SOMA can capture multiple biological dimensions in ccRCCs, which can then serve to further motivate novel applications of the biomarker beyond its initial prescription, including stratifying patient cohorts in a phase II clinical trial based on rPFS to bevacizumab. Although a proof of concept study, these results suggest that radiogenomic SOMAs have the potential to serve as broad powerful tools for noninvasive, multi-scale and multi-feature phenotype analysis which may be of paramount clinical relevance to the design and implementation of clinical trials. Clearly, additional analysis and scaling to other applications and tumour types is warranted, but these results are, nonetheless, promising.

Interestingly, the results of our prospective evaluation of the RRS in phase II trial data suggest that the RRS can stratify patient response (OS and PFS) prior to treatment with bevacizumab and cytoreductive nephrectomy (Fig. 3). The MSKCC criteria are a well-established prognostic tool consisting of clinical and laboratory measures that can be applied across all stages of RCC, similar to the

Table 2 Classification of clinical variables in terms of high- and low-RRS groups. Summary of demographic data, MSKCC and Heng criteria, RECIST response rate and clinical outcomes, as grouped by RRS scores. OS: overall survival. PFS: progression-free survival. * denotes statistically significant differences. Associations with different metastases sites (including the lungs, mediastinum, retroperitneum, adrenal glands, bone, and liver) were assessed and there were no statistically significant correlations when grouped according to the high- and low-RRS groups. The summary table shows the average number of metastatic sites for each RRS group

	RRS (low) (n=14)	RRS (high) (n=27)	*P*-value
Age (median)	62	58	0.47
Gender			0.70
male	9 (64 %)	19 (70 %)	
female	5 (36 %)	8 (30 %)	
Mean Number of Total Sites	1.36	1.81	0.12
adrenal metastasis	0.29	0.22	
bone metastasis	0.14	0.11	
liver metastasis	0.07	0.11	
lung metastasis	0.57	0.70	
mediastinal metastasis	0.14	0.41	
retroperitoneal metastasis	0.14	0.07	
MSKCC			0.15
intermediate	12 (86 %)	15 (56 %)	
poor	2 (14 %)	12 (44 %)	
Heng			0.13
intermediate	4 (29 %)	4 (15 %)	
poor	10 (71 %)	23 (85 %)	
Best Response			0.05*
sd	12 (86 %)	15 (56 %)	
pd	1 (7 %)	8 (30 %)	
pr	1 (7 %)	3 (11 %)	
RRS PFS (median)	>25 months	6 months	0.005*
RRS OS (median)	>37 months	25 months	0.03*

RRS. Although trending towards significance, the MSKCC criteria do not appear to be predictive of bevacizumab response in this cohort, consistent with previous reports [16]. Similarly, the Heng criteria, which were originally derived from a retrospective analysis of a large pool of advanced/mRCC patients treated with a variety of anti-angiogenesis therapies (bevacizumab, sunitinib, or sorafenib) also were not significant in our population. These findings are consistent with our Cox analysis which showed that the RRS was the only significant predictor of PFS when comparing these three prognostic tools. Interestingly, although the trial study population consisted of patients with metastatic disease, stratification of rPFS was solely based on RRS determination (which itself, is based on radiogenomic evaluation of the primary tumour), and was independent of metastatic site involvement and primary tumour size. Finally, we observed that the high-risk score group was more highly enriched in patients whose best response was progressive disease relative to the low-risk group. This could suggest that the RRS is able to identify those patients more likely to progress to early failure prior to initiation of treatment, thus potentially serving as an early screening indicator of likely early progression before progression is determined by RECIST criteria.

Table 3 Multivariate Cox regression analysis of progression-free survival. *denotes statistically significant differences

	Hazard Ratio (exponent)	*P*-value
RRS	8.58	0.04*
MSKCC	1.80	0.34
Heng	1.11	0.90

This study has several limitations. First, a larger sample size would have been ideal for the phase II evaluation of the RRS. Interestingly, the sample size used in our study is consistent with other phase II solid tumour investigator-initiated clinical trials; that this approach was able to discern a significant signal is encouraging. Second, although phase II success rates have dropped below 20 % [30] (hence the need for better means of optimizing trials through improved upfront patient stratification), confirmation of the RRS ability to stratify rPFS in an independent phase III data set would have been ideal. Additionally, although the RRS demonstrates an ability to stratify metastatic RCC patients into groups, more or less likely to derive a rPFS in the setting of pre-surgical bevacizumab independent of other factors, with the provided data, it is not possible for us to distinguish whether the RRS is truly predictive or prognostic. Finally, in the setting of new mandates for low-dose CT, interpretation of some of the features may become more challenging. Future work may need to address this issue in determining the requisite dose settings that permit reproducible interpretation of particular features.

In summary, our results demonstrate that the ccRCC-specific RRS, is a robust, semi-quantitative, radiogenomic-based SOMA that may have properties that enable stratification of patients more or less likely to derive a rPFS in the setting of patients with metastatic RCC undergoing treatment with bevacizumab and cytoreductive nephrectomy. Adaptive clinical trials have shown that methodical coupling of molecular biomarkers with targeted therapeutics can potentially result in the rapid identification of potential candidate therapies for a target population with a high probability of success based on a smaller starting patient sample size [31, 32]. By accounting for

more of the rich information contained in the imaging appearance and characteristics of tumours, we demonstrate that SOMAs have the potential to improve the imaging assessment of tumour response through incorporation of genomic data. These assays may ultimately help to facilitate the widespread translation of clinical molecular assays in a safe, portable, and efficient manner across a wide array of potential applications in the near future.

Acknowledgments MDK is the scientific guarantor of this publication. The authors of this manuscript declare no relationships with any companies whose products or services may be related to the subject matter of the article. The authors state that this work has not received any funding. MZ, NJ, and MDK have significant statistical expertise. Institutional review board approval was obtained from Umea Hospital and the MD Anderson Cancer Center. Written informed consent was obtained from all subjects in this study. None of the study cohort imaging findings or results have been previously reported in the literature. This is a retrospective, diagnostic/prognostic, single institution study.

Compliance with ethical standards

Conflict of interest The authors disclose no potential conflicts of interest.

References

1. Jemal A, Siegel R, Ward E, Murray T, Xu J, Thun MJ (2007) Cancer statistics, 2007. CA Cancer J Clin 57:43–66
2. Kuo MD, Gollub J, Sirlin CB, Ooi C, Chen X (2007) Radiogenomic analysis to identify imaging phenotypes associated with drug response gene expression programs in hepatocellular carcinoma. J Vasc Interv Radiol 18:821–831
3. Kuo MD, Jamshidi N (2014) Behind the numbers: decoding molecular phenotypes with radiogenomics–guiding principles and technical considerations. Radiology 270:320–325
4. Banerjee S, Wang DS, Kim HJ et al (2015) A computed tomography radiogenomic biomarker predicts microvascular invasion and clinical outcomes in hepatocellular carcinoma. Hepatology. doi:10.1002/hep.27877
5. Diehn M, Nardini C, Wang DS et al (2008) Identification of noninvasive imaging surrogates for brain tumor gene-expression modules. Proc Natl Acad Sci U S A 105:5213–5218
6. Jamshidi N, Diehn M, Bredel M, Kuo MD (2014) Illuminating radiogenomic characteristics of glioblastoma multiforme through integration of MR imaging, messenger RNA expression, and DNA copy number variation. Radiology 270:1–2
7. Segal E, Sirlin CB, Ooi C et al (2007) Decoding global gene expression programs in liver cancer by noninvasive imaging. Nat Biotechnol 25:675–680
8. Yamamoto S, Han W, Kim Y et al (2015) Breast cancer: radiogenomic biomarker reveals associations among dynamic contrast-enhanced mr imaging, long noncoding RNA, and metastasis. Radiology. doi:10.1148/radiol.15142698:142698
9. Yamamoto S, Korn RL, Oklu R et al (2014) ALK molecular phenotype in non-small cell lung cancer: CT radiogenomic characterization. Radiology 272:568–576
10. Gevaert O, Xu J, Hoang CD et al (2012) Non-small cell lung cancer: identifying prognostic imaging biomarkers by leveraging public gene expression microarray data–methods and preliminary results. Radiology 264:387–396
11. Rizzo S, Petrella F, Buscarino V et al (2015) CT radiogenomic characterization of EGFR, K-RAS, and ALK mutations in non-small cell lung cancer. Eur Radiol. doi:10.1007/s00330-015-3814-0
12. Klein EA, Cooperberg MR, Magi-Galluzzi C et al (2014) A 17-gene assay to predict prostate cancer aggressiveness in the context of Gleason grade heterogeneity, tumor multifocality, and biopsy undersampling. Eur Urol 66:550–560
13. Sestak I, Dowsett M, Zabaglo L et al (2013) Factors predicting late recurrence for estrogen receptor-positive breast cancer. J Natl Cancer Inst 105:1504–1511
14. Solomon BJ, Mok T, Kim DW et al (2014) First-line crizotinib versus chemotherapy in ALK-positive lung cancer. N Engl J Med 371:2167–2177
15. Jamshidi N, Jonasch E, Zapala M et al (2015) The radiogenomic risk score: construction of a prognostic quantitative, noninvasive image-based molecular assay for renal cell carcinoma. Radiology 277:114–123
16. Jonasch E, Wood CG, Matin SF et al (2009) Phase II presurgical feasibility study of bevacizumab in untreated patients with metastatic renal cell carcinoma. J Clin Oncol 27:4076–4081
17. Sun M, Shariat SF, Cheng C et al (2011) Prognostic factors and predictive models in renal cell carcinoma: a contemporary review. Eur Urol 60:644–661
18. Heng DY, Xie W, Regan MM et al (2013) External validation and comparison with other models of the International Metastatic Renal-Cell Carcinoma Database Consortium prognostic model: a population-based study. Lancet Oncol 14:141–148
19. Heng DY, Xie W, Regan MM et al (2009) 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 27:5794–5799
20. Motzer RJ, Escudier B, Bukowski R et al (2013) Prognostic factors for survival in 1059 patients treated with sunitinib for metastatic renal cell carcinoma. Br J Cancer 108:2470–2477
21. Zeeberg BR, Feng W, Wang G et al (2003) GoMiner: a resource for biological interpretation of genomic and proteomic data. Genome Biol 4:R28
22. Breslow N, Crowley J (1984) A large sample study of the life table and product limit estimates under random censorship. Ann Stat 2:437–453
23. Viera AJ, Garrett JM (2005) Understanding interobserver agreement: the kappa statistic. Fam Med 37:360–363
24. Mekhail TM, Abou-Jawde RM, Boumerhi G et al (2005) Validation and extension of the Memorial Sloan-Kettering prognostic factors model for survival in patients with previously untreated metastatic renal cell carcinoma. J Clin Oncol 23:832–841
25. Finn RS, Crown JP, Lang I et al (2015) The cyclin-dependent kinase 4/6 inhibitor palbociclib in combination with letrozole versus letrozole alone as first-line treatment of oestrogen receptor-positive, HER2-negative, advanced breast cancer (PALOMA-1/TRIO-18): a randomised phase 2 study. Lancet Oncol 16:25–35
26. Robert C, Long GV, Brady B et al (2015) Nivolumab in previously untreated melanoma without BRAF mutation. N Engl J Med 372:320–330
27. Lamuraglia M, Raslan S, Elaidi R et al (2015) mTOR-inhibitor treatment of metastatic renal cell carcinoma: contribution of Choi and modified Choi criteria assessed in 2D or 3D to evaluate tumor response. Eur Radiol. doi:10.1007/s00330-015-3828-7
28. Gerlinger M, Rowan AJ, Horswell S et al (2012) Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med 366:883–892

29. Lindstrom LS, Karlsson E, Wilking UM et al (2012) Clinically used breast cancer markers such as estrogen receptor, progesterone receptor, and human epidermal growth factor receptor 2 are unstable throughout tumor progression. J Clin Oncol 30:2601–2608
30. Arrowsmith J (2011) Trial watch: phase II failures: 2008-2010. Nat Rev Drug Discov 10:328–329
31. Coffey CS, Levin B, Clark C et al (2012) Overview, hurdles, and future work in adaptive designs: perspectives from a National Institutes of Health-funded workshop. Clin Trials 9:671–680
32. Ledford H (2013) 'Master protocol' aims to revamp cancer trials. Nature 498:146–147