




Adv Ther  
<https://doi.org/10.1007/s12325-019-01167-2>

ORIGINAL RESEARCH

# Efficacy and Safety of Approved First-Line Tyrosine Kinase Inhibitor Treatments in Metastatic Renal Cell Carcinoma: A Network Meta-Analysis

Kirsi M. Manz · Klaus Fenchel · Andreas Eilers · Jonathan Morgan ·  
Kirsten Wittling · Wolfram C. M. Dempke 

Received: September 30, 2019  
© The Author(s) 2019

## ABSTRACT

**Introduction:** This network meta-analysis aims to deliver an up-to-date, comprehensive efficacy and toxicity comparison of the approved first-line tyrosine kinase inhibitors (TKIs) for metastatic renal cell carcinoma (mRCC) in order to provide support for evidence-based treatment decisions. Previous NMAs of first-line mRCC treatments either predate the approval of all the first-line TKIs currently available or do not

**Enhanced Digital Features** To view enhanced digital features for this article go to <https://doi.org/10.6084/m9.figshare.11110208>.

**Electronic Supplementary Material** The online version of this article (<https://doi.org/10.1007/s12325-019-01167-2>) contains supplementary material, which is available to authorized users.

K. M. Manz  
Institut für Medizinische Informationsverarbeitung,  
Biometrie und Epidemiologie (IBE), LMU Munich,  
Munich, Germany

K. Fenchel  
Department of Haematology and Oncology,  
Medical School Hamburg (MSH), Hamburg,  
Germany

A. Eilers · J. Morgan · K. Wittling  
EUSA Pharma Ltd, Hemel Hempstead, UK

W. C. M. Dempke (✉)  
Department of Haematology and Oncology,  
University Clinic, LMU Munich, Munich, Germany  
e-mail: wolfram.dempke@web.de

include evaluation of safety data for all treatments.

**Methods:** We performed a systematic literature review and network meta-analysis of phase II/III randomised controlled trials (RCTs) assessing approved first-line TKI therapies for mRCC. A random effects model with a frequentist approach was computed for progression-free survival (PFS) data and for the proportion of patients experiencing a maximum of grade 3 or 4 adverse events (AEs).

**Results:** The network meta-analysis of PFS demonstrated no significant differences between cabozantinib and either sunitinib (50 mg 4/2), pazopanib or tivozanib. The network meta-analysis indicated that in terms of grade 3 and 4 AEs, tivozanib had the most favourable safety profile and was associated with significantly less risk of toxicity than the other TKIs.

**Conclusion:** These network meta-analysis data demonstrate that cabozantinib, sunitinib, pazopanib and tivozanib do not significantly differ in their efficacy, but tivozanib is associated with a more favourable safety profile in terms of grade 3 or 4 toxicities. Consequently, the relative toxicity of these first-line TKIs may play a more significant role than efficacy comparisons in treatment decisions and in planning future RCTs.

**Keywords:** Adverse event; Metastatic renal cell carcinoma; Network meta-analysis; Progression-free survival; Randomised controlled trials; Tyrosine kinase inhibitors

### Key Summary Points

#### Why carry out this study?

A network meta-analysis was performed, evaluating approved first-line tyrosine kinase inhibitors (TKIs) for metastatic renal cell carcinoma (mRCC).

This provided an up-to-date, comprehensive analysis of phase II/III randomized controlled trial data.

#### What was learned from this study?

No significant efficacy differences between approved first-line TKIs were observed.

Tivozanib ranked the most favourable in the analysis of grade 3 and 4 adverse events.

This produced indirect evidence to support clinical decisions and planning of future trials.

## INTRODUCTION

The treatment of metastatic renal cell carcinoma (mRCC) has evolved over the last decade and now harbours several different drug classes [cytokines, tyrosine kinase inhibitors (TKIs), mammalian target of rapamycin inhibitors and immune checkpoint inhibitors (ICIs, IOs)] [1, 2]. European mRCC treatment guidelines [European Association of Urology (EAU) and European Society for Medical Oncology (ESMO)] were updated in 2018/2019, with TKIs recommended as the standard for treating favourable International Metastatic RCC Database Consortium (IMDC) risk-group patients, but as optional first-line treatments (secondary to ICIs as standard) in intermediate- or poor-risk

patients [1, 2]. Guidelines now include the following TKIs as first line: cabozantinib, pazopanib, sorafenib, sunitinib and tivozanib [1, 2]. The first-line TKI options have varying potency and selectivity for vascular endothelial growth factor receptors (VEGFRs) [3, 4]. Pazopanib, sorafenib, sunitinib and cabozantinib are considered multi-targeted TKIs because they inhibit several tyrosine kinases, such as platelet-derived growth factor receptor and c-KIT, in addition to VEGFR, whereas tivozanib has been shown to potently and selectively target all three VEGF receptors [3, 4]. All approved first-line TKIs have demonstrated anti-tumour activity, but it is proposed that the off-target effects contribute to differences between the toxicity profiles [1–4]. Examples of off-target toxicities include diarrhoea, fatigue and hand-foot syndrome, whereas VEGF-associated toxicities include hypertension and hypothyroidism [3, 4]. With many TKIs available, and given the lack of head-to-head randomised controlled trials (RCTs), it is important to evaluate the relative efficacy and toxicity of each TKI to support an evidence-based approach to treatment. Two previous network meta-analyses (NMAs) of first-line mRCC treatments have been carried out: one predates the approval of cabozantinib as a first-line treatment; the other demonstrated that there was no significant difference in progression-free survival (PFS) associated with several first-line treatments, but this analysis did not include safety profile data for all therapies studied [5, 6]. Toxicity may be an important differentiator between these treatments. This NMA aims to provide an up-to-date, comprehensive efficacy and toxicity comparison between each of the approved first-line TKIs for mRCC.

## METHODS

### Systematic Literature Review Methodology

We performed a systematic literature review and NMA of phase II/III RCTs assessing approved first-line TKI therapies for mRCC. PubMed, ClinicalTrials.gov, Embase, Medline, the

Cochrane Central Register of Controlled Trials, Web of Science and conference abstracts from the American Society of Clinical Oncology (ASCO), ESMO and ASCO-Genitourinary were searched independently by two authors (WD and AE). Only English language publications from database inception to 15 January 2019 were included. Search terms included: randomised clinical trial; metastatic renal cell carcinoma; advanced renal cell carcinoma; tyrosine kinase inhibitor; first-line; phase II trial; phase III trial; immunotherapy; progression-free survival; adverse events. Results were restricted to phase II and phase III RCTs. Bibliographies of review articles and editorials were manually searched. The literature review process followed Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [7]: two authors (JM and KF) independently evaluated data from eligible studies, which were then checked by a third author (KW), and any disagreements were resolved by discussions moderated by a fourth author (WD). A bias risk assessment was conducted using The Cochrane Collaboration's tool [8]. All RCTs comparing a first-line TKI with other TKIs, placebo or interferon- $\alpha$  (a historic standard of care) were included; other comparators were excluded. Additional exclusion criteria included: non-randomised trials, retrospective studies, second-line or later-line studies, case reports and TKIs not approved for first-line therapy.

The outcome measure used to evaluate efficacy in the NMA was PFS, as first reported by the study authors (i.e., either independent review committee or investigator assessed). In some instances, time-to-progression (TTP) data were substituted as a close approximate. The toxicity outcome measure used was the proportion of patients experiencing a maximum of grade 3 or 4 adverse events (AEs). Contact with the study authors was attempted (email and/or telephone) to obtain missing information. In some cases it was possible to derive missing information from available data using formulae adapted from Woods et al. (2010) and Altman and Bland (2011) [9, 10].

Data collection for this work is based on previously conducted studies and does not

contain any study with human participants or animals performed by any of the authors. The data are solely obtained from published studies.

## Statistical Analysis

All statistical analyses were performed using R software and the netmeta package [11, 12]. For the network analyses, the treatment effect [log (PFS hazard ratio)] and toxicity effect [log (relative risk of % patients experience grade 3 or 4 AE)], and estimates of the standard error of each, were inputted into the model, using the data are collated in Table 1. Network diagrams were produced: the thickness of the connecting lines represents the strength of evidence for a treatment effect [11–13].

We performed an NMA using a random effects model with a frequentist approach [14, 15]. A fixed effect model was tested but significant heterogeneity was detected in the overall network ( $Q_{\text{total}} = 13.99, p = 0.0073$ ), which could be decomposed into considerable heterogeneity between the designs ( $Q_{\text{between}} = 12.00, p = 0.0074$ ) and non-significant heterogeneity within the designs ( $Q_{\text{within}} = 1.99, p = 0.1586$ ). A design consists of a pairwise comparison of two treatments such as sorafenib vs. tivozanib. Therefore, a random effects model was selected over a fixed-effects model to account for this potential heterogeneity (different study designs, populations, treatment arms, etc.) [16].

Sources of inconsistency in the random effects model were investigated by the generation of net heat plots: the colour-scale shading of each box indicates whether the design was a source of inconsistency (red) or supported other evidence (blue), and the area of the grey boxes indicates the contribution of direct-comparison evidence and indirect evidence (given in the columns) to the network estimate of a comparison (shown in the rows) [13].

Treatments were ranked by calculating  $P$  scores using the netrank function of the netmeta package [11, 17].  $P$  scores measure the extent of certainty that a treatment is better than another treatment, averaged over all competing treatments, while taking the precision into account [17].

Table 1 RCTs included in the network meta-analysis

Study	Treatment 1	Treatment 2	$N_{ITT}$ $n_1$ $n_2$ safety population in treatment 1 and 2	mPFS treatment 1 vs. 2	HR treatment 1 vs. 2 (95% CI)	AE treatment 1 vs. 2	Prognostic risk group inclusion criteria	Study phase
Motzer et al. 2007 [19]	Sunitinib (50 mg qd 4/2)	Interferon-alfa (9 MU thrice weekly)	$N$ 750 $n_1$ 375 $n_2$ 360	11 vs. 5	0.42 (0.32–0.54)	0.48 vs. 0.17	NA	Phase III
Escudier et al. 2009 [21]	Interferon-alfa (9 MU thrice weekly)	Sorafenib (400 mg bid)	$N$ 189 $n_1$ 90 $n_2$ 97	5.6 vs. 5.7	0.88 (0.61–1.27)	0.554 vs. 0.701	NA	Phase II
Sternberg et al. 2010 [18]	Pazopanib (800 mg qd)	Placebo	$N$ 233 $n_1$ 155 $n_2$ 78	11.1 vs. 2.8	0.40 (0.27–0.60)	0.40 vs. 0.20	NA	Phase III
Motzer et al. 2011 [29]	Sunitinib (50 mg qd 4/2)	Sunitinib (37.5 mg qd CDD)	$N$ 292 $n_1$ 144 $n_2$ 145	9.9 vs. 7.1 <sup>a</sup>	0.773 (0.572–1.044) <sup>a</sup>	0.56 vs. 0.58	NA [30]	Phase II
Nosov et al. 2012 [20]	Tivozanib (1.5 mg qd 3/1)	Placebo	$N$ 118 $n_1$ 61 $n_2$ 57	10.3 vs. 3.3	0.45 (0.245–0.826) <sup>d</sup>	0.13 vs. 0.08	NA	Phase II
Motzer et al. 2013a [22]	Tivozanib (1.5 mg qd 3/1)	Sorafenib (400 mg bid)	$N$ 517 $n_1$ 259 $n_2$ 257	11.9 vs. 9.1	0.797 (0.639–0.993)	0.25 <sup>b</sup> vs. 0.70	NA	Phase III
Motzer et al. 2013b [24]	Pazopanib (800 mg qd)	Sunitinib (50 mg qd 4/2)	$N$ 1110 $n_1$ 554 $n_2$ 548	8.4 vs. 9.5	1.05 (0.90–1.22)	0.54 vs. 0.69	NA	Phase III

Table 1 continued

Study	Treatment 1	Treatment 2	$N_{ITT}$ $n_1$ $n_2$ safety population in treatment 1 and 2	mPFS treatment 1 vs. 2	HR treatment 1 vs. 2 (95% CI)	AE treatment 1 vs. 2	Prognostic risk group inclusion criteria	Study phase
Tomita et al. 2014 [28]	Sunitinib (50 mg qd 4/2)	Sorafenib (400 mg bid)	$N$ 124 $n_1$ 57 $n_2$ 63	8.7 vs. 7.0	0.67 (0.42–1.08)	0.70 vs. 0.71	MSKCC risk: favourable or intermediate [39]	Phase III
Lee et al. 2015 [25]	Sunitinib (50 mg qd 4/2)	Sunitinib (50 mg qd 2/1)	$N$ 74 $n_1$ 36 $n_2$ 38	10.1 vs. 12.1 <sup>a</sup>	0.75 (0.44–1.27) <sup>a</sup>	0.72 vs. 0.53	NA	Phase II
Eichelberg et al. 2015 [26]	Sorafenib (400 mg bid)	Sunitinib (50 mg qd 4/2)	$N$ 365 $n_1$ 177 $n_2$ 176	12.5 vs. 14.9 <sup>c</sup>	1.01 <sup>c</sup> (0.81–1.27) <sup>c</sup>	0.66 vs. 0.67	MSKCC risk: favourable or intermediate	Phase III
Choueiri et al. 2017 [23]	Cabozantinib (60 mg qd)	Sunitinib (50 mg bid)	$N$ 157 $n_1$ 78 $n_2$ 72	8.2 vs. 5.6	0.66 (0.46–0.95)	0.67 vs. 0.68	IMDC risk: intermediate or poor	Phase II

Table 1 continued

Study	Treatment 1	Treatment 2	$N_{ITT}$ $n_1$ $n_2$ safety population in treatment 1 and 2	mPFS treatment 1 vs. 2	HR treatment 1 vs. 2 (95% CI)	AE treatment 1 vs. 2	Prognostic risk group inclusion criteria	Study phase
Retz et al. 2019 [27]	Sorafenib (400 mg bid)	Pazopanib (800 mg qd)	$N$ 377 $n_1$ 183 $n_2$ 183	8.6 vs. 12.9 <sup>c</sup>	1.36 <sup>c</sup> (1.11–1.68)	0.59 vs. 0.64	MSKCC risk: favourable or intermediate	Phase III

The data used to compute the network meta-analysis were the given PFS hazard ratio (HR) and 95% confidence intervals (95% CI) and the proportion of patients experiencing grade 3 or 4 AEs in each treatment arm. Certain RCT characteristics are also provided

#/#: a cycle of a number of weeks on a drug followed by the number of weeks off drug

AE adverse event, *bid* twice daily, *CDD* continuous daily dose, *CI* confidence interval, *ECOG* Eastern Cooperative Oncology Group, *HR* hazard ratio, *IMDC* International Metastatic Renal Cell Carcinoma Database Consortium, *ITT* intention to treat, *mPFS* median progression-free survival, *MSKCC* Memorial Sloan Kettering Cancer Center, *PFS* progression-free survival, *qd* once daily, *TTP* time to progression

<sup>a</sup> TTP data were used. PFS data were not available for Lee et al. [25]. PFS data were available for Motzer et al. (2011/2012); the PFS HR (HR 0.77, 95% CI 0.58–1.02) was very similar to the TTP HR [29, 30]

<sup>b</sup> Updated safety data from the internal tivozanib safety data bank (data on file)

<sup>c</sup> Total PFS data for the sequence of the two treatments were used to account for patients who switched treatment because of an AE prior to progression or death on first-line treatment. Without complete information on data-censoring practices in each study, this total PFS primary end point was considered comparable to the primary end point PFS values used in other studies

<sup>d</sup> 95% CI value was calculated from the given *p* value of *p* = 0.01

<sup>e</sup> Represents 90% CI, which was converted into 95% CI for the data analysis

## RESULTS

The systematic literature search identified 699 unique references, which, after review, revealed 12 that fitted the screening criteria (Fig. 1). Table 1 presents the RCT data input into the model and some key trial characteristics. The RCTs directly comparing TKIs to either IFN- $\alpha$  or placebo demonstrated significant improvements in PFS [18–20], except for sorafenib versus IFN- $\alpha$  [21]. RCTs directly comparing TKIs to one another demonstrated mixed results: some demonstrated significant improvements in PFS [22, 23] or established non-inferiority [24] while others did not [25–30]. As data from two studies were available in abstract form only, we were unable to access their risk of bias. All other studies included were open-label trials. We felt that all studies were at low risk of attrition and reporting bias.

For the efficacy NMA, we included all 12 studies with a total of 4306 patients (Fig. 2a), and for the safety analysis, we included data for all 12 studies with a total of 4243 patients (Fig. 3a). The strength of evidence for the sunitinib (50 mg 2/1) dosing regimen was the weakest in the NMA (Fig. 2a, 3a), perhaps because of the small sample size (Table 1). The NMA output data are tabulated in Appendix Tables S1 and S2. The eligibility criteria for the 12 studies varied, for example the majority of studies did not specify a Memorial Sloan Kettering Cancer Center (MSKCC) prognostic group as an entry criteria, except for three studies that only enrolled patients with a favourable or intermediate MSKCC risk score [26–28] and one that enrolled patients of intermediate or poor IMDC risk category [23]. These differences in eligibility criteria can be a potential source of heterogeneity, which is partially accounted for in the random effects

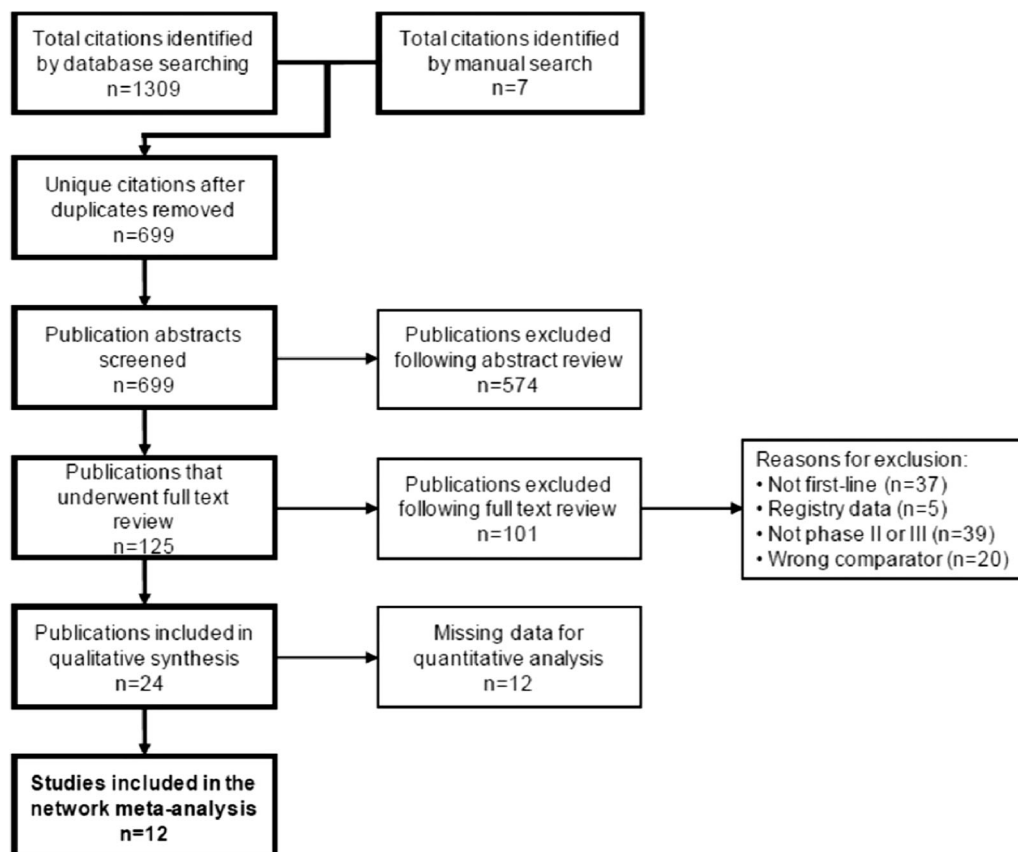
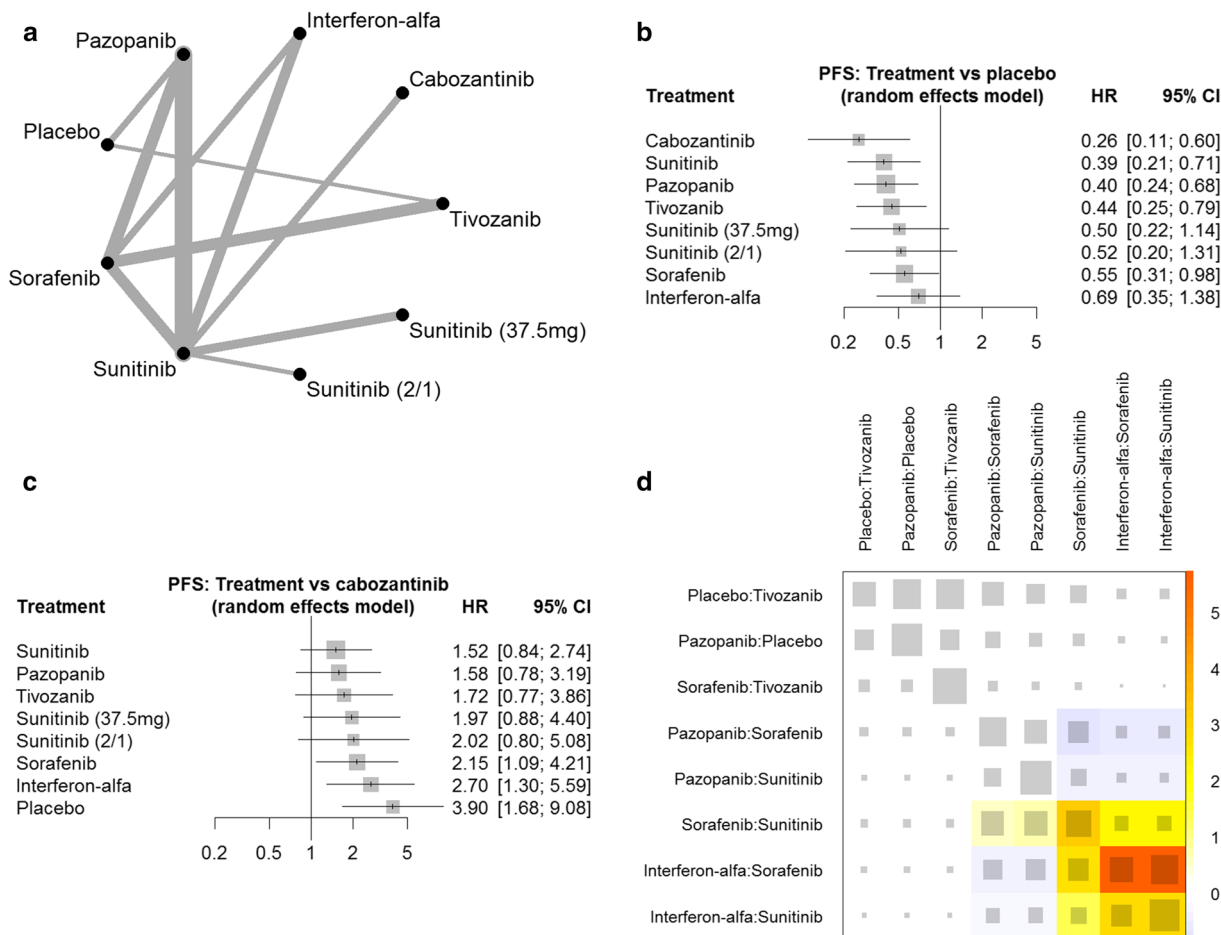


Fig. 1 PRISMA diagram



**Fig. 2** Network meta-analysis of PFS: **a** network diagram; **b** forest plot, with placebo as the comparator. *PFS* progression-free survival, *HR* hazard ratio, *CI* confidence interval; **c** forest plot, with cabozantinib as the comparator; **d** net heat plot. Red areas are related to inconsistency,

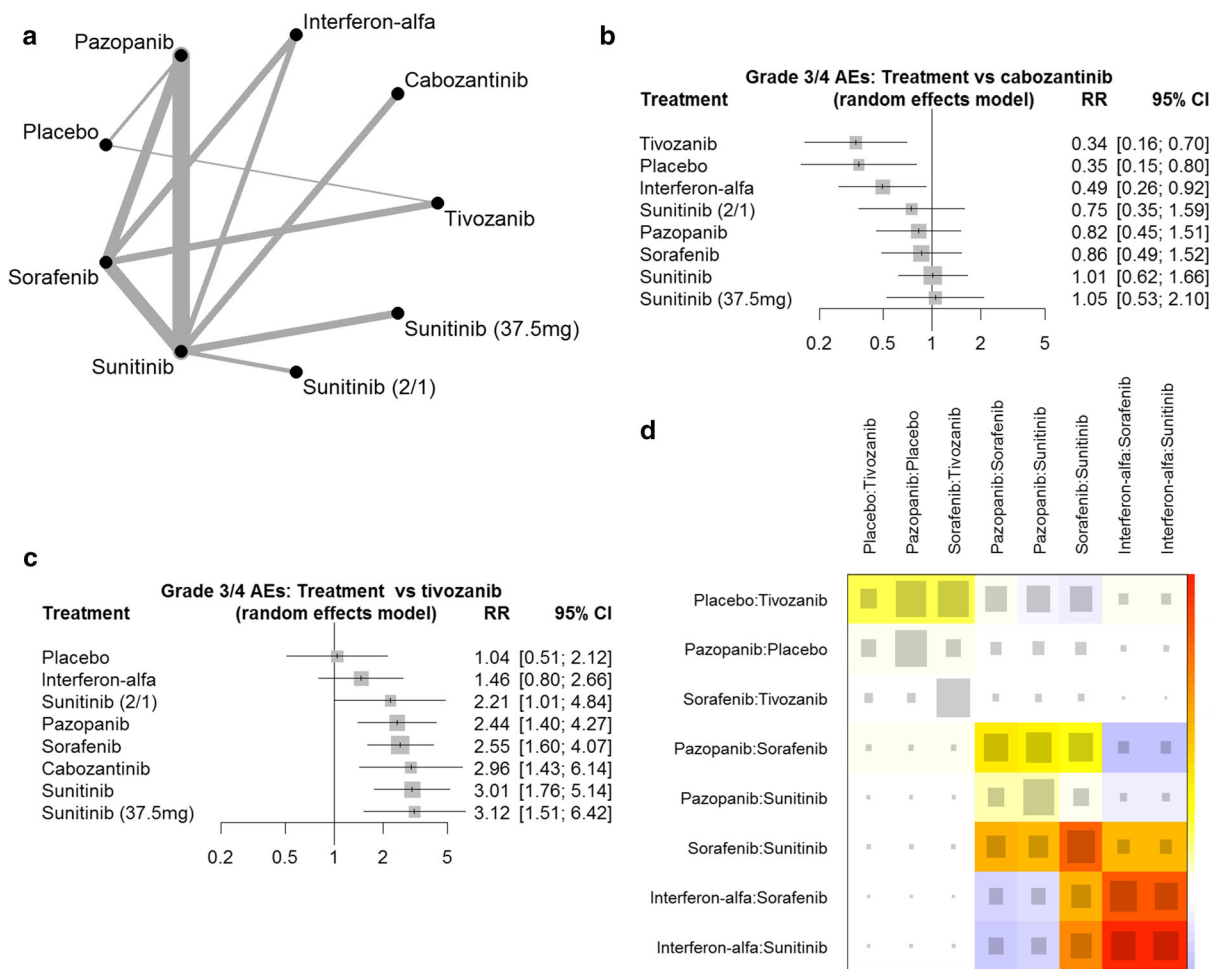
whereas blue areas support other evidence gained from the network. The area of a grey box indicates the contribution of the direct estimate of the pairwise comparison in the column to a network estimate in a row

model. When analysing for specific sources of heterogeneity, the studies including only favourable or intermediate MSKCC risk patients [26–28] were not collectively found to be a significant cause of inconsistency (Figs. 2d, 3d). The net heat plots also show that these studies contribute important indirect evidence to the model. It was not possible to analyse the effect of restrictive MSKCC eligibility criteria at the other end of the prognostic risk spectrum in this way because only one cabozantinib study that used these criteria was included [23].

Figure 2b shows the NMA results of the indirect efficacy comparison with placebo; the

confidence intervals demonstrate that cabozantinib, sunitinib [standard regimen (50 mg 4/2)], pazopanib, tivozanib and sorafenib treatments were significantly different from placebo, whereas the alternative sunitinib dosing regimens [50 mg 2/1 and 37.5 mg continuous daily dose (CDD)] were not. Cabozantinib had the highest probability of being the best treatment in terms of PFS (*P* score 0.9481), followed by sunitinib, pazopanib and tivozanib (*P* score 0.7411, 0.6914 and 0.5988, respectively) (Table 2). When treatments were indirectly compared with cabozantinib, it was clear that there was no significant difference in PFS





**Fig. 3** Network meta-analysis of percentage grade 3 or 4 AEs: **a** network diagram; **b** forest plot, with cabozantinib as the comparator. *AE* adverse event, *RR* relative risk, *CI* confidence interval; **c** forest plot, with tivozanib as the comparator; **d** net heat plot. Red areas are related to

inconsistency, whereas blue areas support other evidence gained from the network. The area of a grey box indicates the contribution of the direct estimate of the pairwise comparison in the column to a network estimate in a row

between the first-line TKIs, with the exception of sorafenib, which was associated with a significantly shorter PFS (Fig. 2c). Figure 3b shows the NMA toxicity results for the indirect comparison of first-line TKIs with cabozantinib: the confidence intervals demonstrate that tivozanib, placebo and IFN- $\alpha$  were associated with a significantly lower likelihood of grade 3/4 AEs. Calculation of *P* scores confirms that tivozanib has a 92.6% probability of having the least toxicity (Table 2). Indirect toxicity comparison with tivozanib demonstrates that the grade 3/4 safety profile of tivozanib is significantly different from all other first-line TKIs (Fig. 3).

## DISCUSSION

The therapeutic arsenal at hand to treat kidney cancer patients is evolving rapidly with novel IO-IO combinations (e.g., nivolumab plus ipilimumab—CheckMate-214 trial [33]) or IO-TKI combinations (e.g., avelumab plus axitinib—Javelin-101 trial [34]; pembrolizumab plus axitinib—Keynote-426 trial [35]) now being approved by the FDA and/or EMA as first-line combination treatment strategies. These combinations have improved treatment outcome dramatically with a significant overall survival

**Table 2** Network meta-analysis of PFS (left) and percentage grade 3 or 4 AEs (right): *P* score ranking

Efficacy		Safety	
Treatment	<i>P</i> score	Treatment	<i>P</i> score
Cabozantinib	0.9481	Tivozanib	0.9261
Sunitinib (50 mg 4/2)	0.7411	Placebo	0.9021
Pazopanib	0.6914	Interferon-alfa	0.7659
Tivozanib	0.5988	Sunitinib (50 mg 2/1)	0.4889
Sunitinib (37.5 mg CDD)	0.4714	Pazopanib	0.4254
Sunitinib (50 mg 2/1)	0.4528	Sorafenib	0.3844
Sorafenib	0.3651	Cabozantinib	0.2367
Interferon-alfa	0.1930	Sunitinib (37.5 mg CDD)	0.1926
Placebo	0.0384	Sunitinib (50 mg 4/2)	0.1778

*CDD* continuous daily dose, *PFS* progression-free survival

benefit (Keynote-426 and CheckMate-214). However, the observed clinical benefit in these trials was associated with increased grade 3/4 AEs: 75.8% for pembrolizumab/axitinib [35] and 71.2% for avelumab/axitinib [34]. For the IO-IO combination, grade 3 and 4 AEs were reported to be 46% and 63%, respectively, with a treatment discontinuation rate of 22% due to adverse events [33], suggesting that there is still a role for a single TKI treatment especially in elderly and less fit patients; however, the optimal TKIs for these patients remain unclear.

This NMA was conducted in an attempt to provide a comprehensive comparison of the efficacy and safety of approved first-line TKIs for advanced and metastatic renal cell carcinoma. The efficacy NMA *P* scores ranked the first-line TKIs from highest probability of efficacy downward as follows: cabozantinib, sunitinib (50 mg 4/2); pazopanib and tivozanib, sunitinib (37.5 mg CDD); sunitinib (50 mg 2/1), sorafenib. It is possible that the small sample size influenced the sunitinib (50 mg 2/1) ranking result. Cabozantinib had a 94.8% probability of being the best treatment in terms of PFS; however, several other treatments also had *P* scores > 50%, and the confidence intervals demonstrate that no significant differences between cabozantinib and either sunitinib

(50 mg 4/2), pazopanib or tivozanib were observed. However, it should be noted that this NMA is underpowered with wide 95% CIs and we cannot conclude “similar efficacy” as our NMA was not an equivalence trial, which is a common feature of all NMAs published so far.

Taken together, it is not possible to produce a clear hierarchy of first-line TKIs based on significant differences in efficacy. Consequently, the toxicity of these TKIs may play a more significant role in treatment decisions. The NMA indicated that in terms of grade 3 and 4 AEs, tivozanib had the most favourable safety profile and was shown to be associated with significantly less risk of toxicity than other TKIs. This result was consistent with the high specificity of tivozanib for VEGFR compared with other multikinase inhibitors and the hypothesis that fewer off-target side effects occur [3, 4].

A previous NMA of mRCC treatments did not include safety profile data for all therapies included [5]. To produce a comprehensive analysis of approved first-line TKIs, unpublished missing safety data were sought and obtained from six studies. To this end, an updated value from the internal tivozanib safety data bank for the proportion of patients experiencing grade 3 or 4 AEs was also included (data on file). The NMA was also computed using the Motzer et al.

(2013) grade 3 or 4 AE data [22], and a similar *P* score rank position was held by tivozanib relative to other TKIs, but tivozanib ranked lower than placebo and IFN- $\alpha$ , respectively (data not shown).

A favourable safety profile has the direct benefit to patient quality of life of experiencing fewer side effects and may also be associated with simplified management owing to fewer dose interruptions or dose reductions that are required to mitigate side effects [3]. Indeed, in the tivozanib versus sorafenib RCT, tivozanib was associated with significantly fewer dose reductions and interruptions due to AEs than sorafenib [22]. Furthermore, low toxicity is a key characteristic of a therapy potentially suitable for use in combination therapy. The CheckMate-016 trial demonstrated that the combination of either sunitinib or pazopanib with nivolumab, an anti-PD-1 antibody, resulted in a high incidence of grade 3/4 AEs, making neither combination suitable for use [31]. The combinations of cabozantinib with nivolumab and tivozanib with nivolumab are both under investigation [32, 33]. Among these TKIs, our safety data NMA results predict that tivozanib has the greatest chance of being a suitable partner to nivolumab, although the specific overlap of each drug's safety profile and any drug-drug interaction will also have influence. As new drugs enter the mRCC treatment landscape, there will be even more combination therapy options to pursue [4].

Cabozantinib had the greatest probability of having the highest efficacy in the NMA; however, the cabozantinib RCT was conducted only in intermediate or poor MSKCC risk patients; hence, it is only indicated in these patients [23, 34]. The NMA data support the ESMO and EAU guidelines that include cabozantinib as a TKI option in poor- or intermediate-risk patients [1, 2].

The current EAU guideline recommends that tivozanib is not used in first-line mRCC treatment because the evidence is considered inferior to other recommended TKIs [2]. This NMA addresses this by providing indirect evidence to supplement the direct evidence for tivozanib. Specifically, the net heat plot diagram (Fig. 2d) demonstrates that while the sorafenib-tivozanib

comparison strongly relies on the direct evidence, the placebo-tivozanib network efficacy estimate gains most evidence from indirect comparisons. The NMA results show tivozanib was not associated with significantly worse efficacy compared with other first-line TKIs and demonstrated a clearly reduced incidence of grade 3 or 4 AEs. These data support the ESMO guidelines, which recommend tivozanib as a first-line mRCC treatment in favourable-risk patients, alongside sunitinib and pazopanib, and as a TKI option in intermediate-risk patients, alongside cabozantinib, sunitinib and pazopanib [1].

In addition, the efficacy NMA data showing that sorafenib PFS is significantly different from cabozantinib, and the *P* score ranking that suggests sorafenib is less likely than either alternative sunitinib regimen to be the best treatment, which were themselves shown to be not significantly different from placebo, provide support for the guidelines that (1) no longer recommend sorafenib as a first-line treatment, apart from in limited-choice settings, and (2) suggest that robust data to support the use of alternative sunitinib dose regimens are lacking [1, 2].

Limitations of meta-analyses using aggregate data have been discussed previously [5]. As the confidence intervals in our analysis and other published NMAs [5, 36] are relatively wide, results need to be treated with caution. In our NMW we used a  $p < 0.05$  threshold to judge the statistical significance of our findings, which means that the results are statistically significant if the confidence intervals do not include the value of 1 (for HR and relative risk). The forest plot for cabozantinib (Fig. 2c) with the HR being  $> 1$  is indicative of inferior efficacy of all other treatments compared with cabozantinib. However, after including the uncertainty around the point estimate (i.e., 95% CIs) for the other TKIs, it becomes clear that the "true" effect could be better for cabozantinib (HR  $< 1$ , lower limit of the CI) or worse than cabozantinib (HR  $> 1$ , upper limit of the CI), which then should be considered as being not statistically significant. However, it has to be taken into account that for cabozantinib only one phase II trial (CABSUN Trial,  $N = 157$ ) had been

published [23]. The results of this trial remain controversial and a number of concerns have been raised because of the poor efficacy of sunitinib in the control arm in which the overall survival (OS) was found to be much worse than in the majority of other studies [5]. On the other hand, tivozanib appeared to be the TKI of choice in our NMA; however, the drug is not approved in the USA because of poor OS outcomes [22]. Finally, one can argue that grade 3–4 AEs may not entirely reflect treatment-related toxicity, suggesting that for an appropriate interpretation of any NMA clinical context is required. Although we have demonstrated no statistically significant differences in the efficacy of the TKIs approved as first-line therapy, it does not rule out the possibility that there still might be one when analysing another data set.

Specific to this NMA, it may be argued that the omission of an evaluation of OS is a limitation. In this regard it should be noted that none of the currently approved TKIs for first-line treatment of advanced or metastatic renal cell cancers has shown a significant OS benefit so far, which prompted us to omit an extensive OS analysis. In addition, some other recently published NMAs have provided evidence that no single TKI treatment appeared to be superior to its comparators for objective response rate (ORR) and was not predictive of OS [36].

However, PFS has been shown to be predictive of OS in TKI-treated patients and, although strict surrogacy has not been established, the US Food and Drug Administration has indicated PFS end points are acceptable [37, 38]. Furthermore, OS can be impacted by differences in sequential therapy, evidenced by the fact that none of the studies included in this analysis demonstrated a significant OS benefit over its comparator; several of these studies suggest OS results were possibly confounded by cross-over to second-line therapies, and some studies specifically evaluated cross-over in a switch design [20–27, 30, 39–41]. Another possible limitation is that grade 1 and 2 toxicities were not included in the analysis. By definition, higher grade AEs are more critical; however, using the IFN- $\alpha$  safety profile as an example, which includes mostly grade 1 and 2 AEs (resulting in it ranking higher than all TKIs except

tivozanib by *P* score), side effects such as fatigue are known to be challenging to manage [42]. Finally, the results of our NMA cannot be directly applied to clinical practice because cabozantinib is only approved for use in intermediate or poor IMDC prognostic risk patients. The cabozantinib trial is one of four trials (out of 12) included in the NMA that had restrictive prognostic risk group entry criteria. We included these studies to provide a comprehensive analysis of all first-line approved TKIs because there was an insufficient number of trials to analyse them separately.

## CONCLUSIONS

In this NMA no statistically significant differences in the efficacy among cabozantinib, sunitinib, pazopanib and tivozanib could be detected; however, tivozanib appeared to be associated with a more favourable safety profile in terms of grade 3 or 4 toxicities. The findings of this NMA may bolster information from pairwise comparisons to shape mRCC clinical decision-making and to assist planning of future RCTs.

## ACKNOWLEDGEMENTS

We thank the study authors who provided unpublished data. The authors are indebted to Prof. Timothy Eisen (University of Cambridge, UK) and Prof. Viktor Gruenwald (University of Essen, Germany) for critical reading of the manuscript.

**Funding.** EUSA Pharma (UK) Ltd. funded the development of this manuscript. The journal's Rapid Service Fee and Open Access publication were sponsored by EUSA Pharma Ltd (UK).

**Medical Writing Assistance.** We thank Dr Hannah Jedrey of TVF Communications (London, UK) for her medical writing support. This was funded by EUSA Pharma (UK) Ltd.

**Authorship.** All named authors meet the International Committee of Medical Journal

Editors (ICMJE) criteria for authorship for this article, take responsibility for the integrity of the work as a whole, and have given their approval for this version to be published.

**Prior Presentation.** Data presented in part at ASCO annual meeting, Chicago, IL, 31 May–4 June 2019.

**Disclosures.** Wolfram Dempke, Kirsi Manz and Klaus Fenchel have nothing to disclose. Andreas Eilers is an employee of EUSA Pharma Ltd. (UK). Jonathan Morgan is an employee of EUSA Pharma Ltd. (UK). Kirsten Wittling is an employee of EUSA Pharma Ltd. (UK).

**Compliance with Ethics Guidelines.** This article is based on previously conducted studies and does not contain any study with human participants or animals performed by any of the authors. The data are solely obtained from published studies.

**Data Availability.** All data generated or analysed during this study are included in this published article and as supplementary information files.

**Open Access.** This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

## REFERENCES

- Escudier B, et al. Renal cell carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2019. <https://doi.org/10.1093/annonc/mdz056>.
- Ljungberg B, Albiges L, Bensalah K, Bex A. EAU-RCC-Guidelines-2018-large-text.pdf. 2018. <https://uroweb.org/wp-content/uploads/EAU-RCC-Guidelines-2018-large-text.pdf>. Accessed 11 Apr 2019.
- Wong HH, Eisen T. Tivozanib for the treatment of metastatic renal cancer. *Expert Rev Anticancer Ther*. 2013;13:649–60.
- Bhargava P, Robinson MO. Development of second-generation VEGFR tyrosine kinase inhibitors: current status. *Curr Oncol Rep*. 2011;13:103–11.
- Wallis CJD, Klaassen Z, Bhindi B, et al. First-line systemic therapy for metastatic renal cell carcinoma: a systematic review and network meta-analysis. *Eur Urol*. 2018;74:309–21.
- Chang X, Zhang F, Liu T, et al. Comparative efficacy and safety of first-line treatments in patients with metastatic renal cell cancer: a network meta-analysis based on phase 3 RCTs. *Oncotarget*. 2016;7:15801–10.
- Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Ann Intern Med*. 2009;151:65–94.
- Higgins JPT, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011;343:d5928.
- Woods BS, Hawkins N, Scott DA. Network meta-analysis on the log-hazard scale, combining count and hazard ratio statistics accounting for multi-arm trials: a tutorial. *BMC Med Res Methodol*. 2010;10:54.
- Altman DG, Bland JM. How to obtain the confidence interval from a P value. *BMJ*. 2011;343:d2090.
- Rücker G, Krahn U, König J, Efthimiou U, Schwarzer G. Netmeta: network meta-analysis using frequentist methods. R package version 1.0-1. 2019. <https://CRAN.R-project.org/package=netmeta>.
- Schwarzer G, Carpenter JR, Rücker TG. Meta-analysis with R. New York: Springer International Publishing; 2015.

13. Krahn U, Binder H, König J. A graphical tool for locating inconsistency in network meta-analyses. *BMC Med Res Methodol.* 2013;13:35.
14. Rücker G. Network meta-analysis, electrical networks and graph theory. *Res Synth Methods.* 2012;3:312–24.
15. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials.* 1986;7:177–88.
16. Fleiss JL. The statistical basis of meta-analysis. *Stat Methods Med Res.* 1993;2:121–45.
17. Rücker G, Schwarzer G. Ranking treatments in frequentist network meta-analysis works without resampling methods. *BMC Med Res Methodol.* 2011;15:58.
18. Sternberg CN, Davis ID, Mardiak J, et al. Pazopanib in locally advanced or metastatic renal cell carcinoma: results of a randomized phase III trial. *J Clin Oncol.* 2010;28:1061–8.
19. Motzer RJ, Hutson TE, Tomczak P, et al. Sunitinib versus interferon alfa in metastatic renal-cell carcinoma. *N Engl J Med.* 2007;356:115–24.
20. Nosov DA, Nosov D, Eisen T, et al. Antitumor activity and safety of tivozanib (AV-951) in a phase II randomized discontinuation trial in patients with renal cell carcinoma. *J Clin Oncol.* 2012;30:1678–85.
21. Escudier B, Nosov D, Eisen T, et al. Randomized phase II trial of first-line treatment with sorafenib versus interferon Alfa-2a in patients with metastatic renal cell carcinoma. *J Clin Oncol.* 2009;27:1280–9.
22. Motzer RJ, Nosov D, Eisen T, et al. Tivozanib versus sorafenib as initial targeted therapy for patients with metastatic renal cell carcinoma: results from a phase III trial. *J Clin Oncol.* 2013;31:3791–9.
23. Choueiri TK, Halabi S, Sanford BL, et al. Cabozantinib versus sunitinib as initial targeted therapy for patients with metastatic renal cell carcinoma of poor or intermediate risk: The Alliance A031203 CABOSUN Trial. *J Clin Oncol.* 2017;35:591–7.
24. Motzer RJ, Hutson TE, Cella D, et al. Pazopanib versus sunitinib in metastatic renal-cell carcinoma. *N Engl J Med.* 2013;369:22–31.
25. Lee JL, Kim MK, Park I, et al. Randomized phase II trial of Sunitinib four weeks on and two weeks off versus two weeks on and one week off in metastatic clear-cell type Renal cell carcinoma: RESTORE trial. *Ann Oncol.* 2015;26:2300–5.
26. Eichelberg C, Bedke J, Bögemann M, et al. SWITCH: a randomised, sequential, open-label study to evaluate the efficacy and safety of sorafenib-sunitinib versus sunitinib-sorafenib in the treatment of metastatic renal cell carcinoma. *Eur Urol.* 2015;68:837–47.
27. Retz M, Bedke J, Bögemann M, et al. SWITCH II: phase III randomized, sequential, open-label study to evaluate the efficacy and safety of sorafenib-pazopanib versus pazopanib-sorafenib in the treatment of advanced or metastatic renal cell carcinoma (AUO AN 33/11). *Eur J Cancer.* 1990;107:37–45.
28. Tomita Y, Naito S, Sassa N, et al. Sunitinib versus sorafenib as first-line therapy for patients with metastatic renal cell carcinoma with favorable or intermediate MSKCC risk factors: a multicenter randomized trial, CROSS-J-RCC. *J Clin Oncol.* 2017;32:502–10.
29. Motzer RJ, Hutson TE, Olsen MR, et al. Randomized phase II multicenter study of the efficacy and safety of sunitinib on the 4/2 versus continuous dosing schedule as first-line therapy of metastatic renal cell carcinoma: Renal Effect Trial. *J Clin Oncol.* 2011;29:LBA308.
30. Motzer RJ, Plimack ER, Ernstoff MS, et al. Randomized phase II trial of sunitinib on an intermittent versus continuous dosing schedule as first-line therapy for advanced renal cell carcinoma. *J Clin Oncol.* 2012;30:1371–7.
31. Amin A, Plimack ER, Ernstoff MS, et al. Safety and efficacy of nivolumab in combination with sunitinib or pazopanib in advanced or metastatic renal cell carcinoma: the CheckMate 016 study. *J Immunother Cancer.* 2018;6:109–15.
32. ClinicalTrials.gov. Phase 1/2 Study of tivozanib in combination with nivolumab in subjects with RCC—Full Text View—ClinicalTrials.gov. <https://clinicaltrials.gov/ct2/show/NCT03136627>. Accessed 23 May 2019.
33. Motzer RJ, Rini BI, McDermott DV, et al. Nivolumab plus ipilimumab versus sunitinib in first-line treatment for advanced renal cell carcinoma: extended follow-up of efficacy and safety results from a randomised, controlled, phase 3 trial. *Lancet Oncol.* 2019;20:1370–85.
34. Motzer RJ, Penkov K, Haanen J, et al. Avelumab plus axitinib versus sunitinib for advanced renal-cell carcinoma. *N Engl J Med.* 2019;380:1103–15.
35. Rini BI, Plimack ER, Stus R, et al. Pembrolizumab plus axitinib versus sunitinib for advanced renal-cell carcinoma. *N Engl J Med.* 2019;380:1116–27.
36. Hahn AW, Klaasen Z, Agarwal N, et al. First-line treatment of metastatic renal cell carcinoma: a

- systematic review and network meta-analysis. *Eur Urol Oncol.* 2019. <https://doi.org/10.1016/j.euo.2019.09.002>.
37. Heng DYC, Xie W, Bjarnason GA, et al. Progression-free survival as a predictor of overall survival in metastatic renal cell carcinoma treated with contemporary targeted therapy. *Cancer.* 2011;117:2637–42.
38. Food and Drug Administration. Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics. 2019. <http://regulatory-information/search-fda-guidance-documents/clinical-trial-endpoints-approval-cancer-drugs-and-biologics>. Accessed 23 May 2019.
39. Sternberg CN, Hawkins RE, Wagstaff J, et al. A randomised, double-blind phase III study of pazopanib in patients with advanced and/or metastatic renal cell carcinoma: final overall survival results and safety update. *Eur J Cancer.* 2013;49:1287–96.
40. Motzer RJ, Hutson TE, Tomczak P, et al. Overall survival and updated results for sunitinib compared with interferon alfa in patients with metastatic renal cell carcinoma. *J Clin Oncol.* 2009;27:3584–90.
41. Tomita Y, Fukasawa S, Shinohara N, et al. Sunitinib versus sorafenib as first-line therapy followed by sorafenib and sunitinib for patients with metastatic renal cell carcinoma (RCC) with clear cell histology: a multicenter randomized trial, CROSS-J-RCC. *J Clin Oncol.* 2017;35:469–74.
42. Ravaud A. Treatment-associated adverse event management in the advanced renal cell carcinoma patient treated with targeted therapies. *Oncologist.* 2011;16(Suppl 2):32–44.