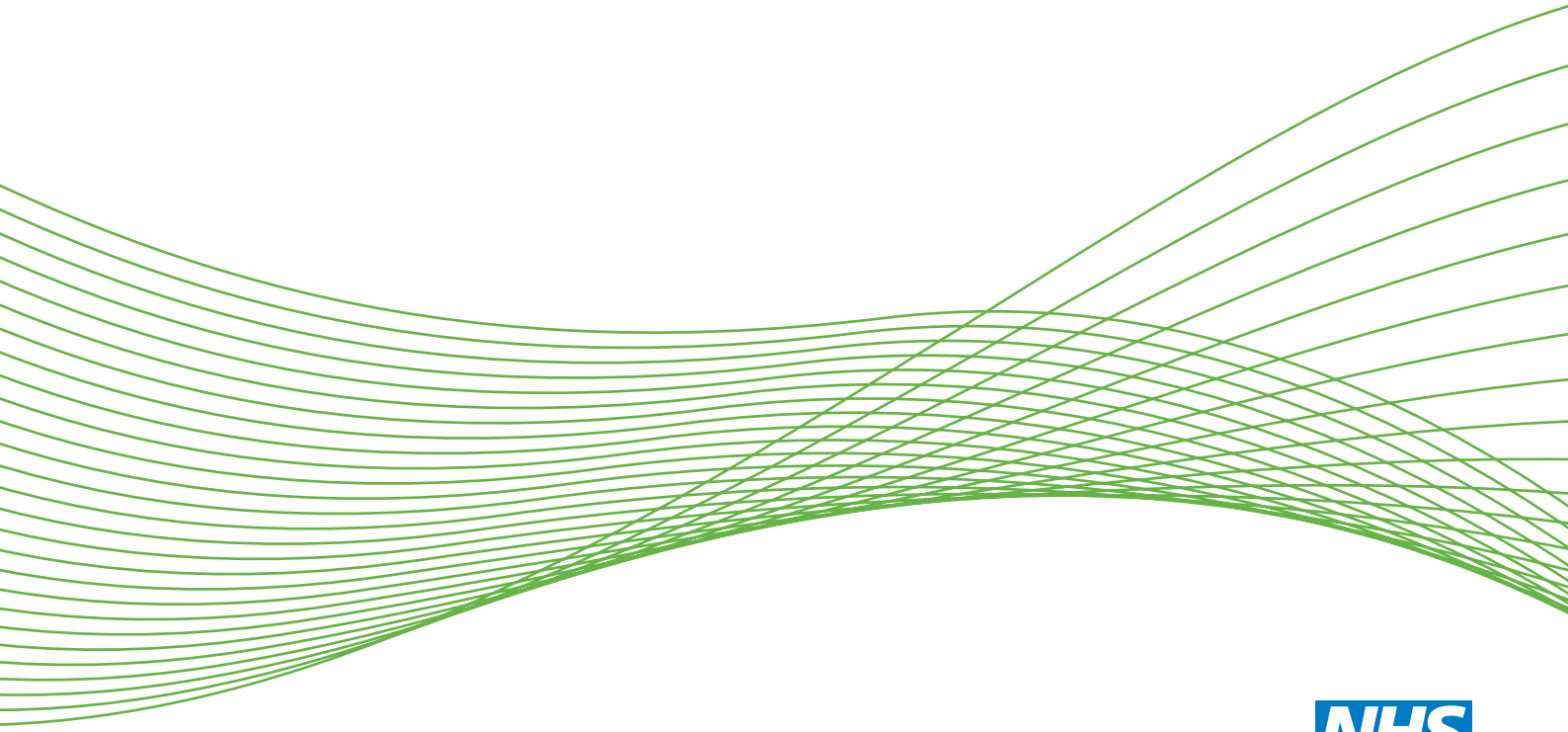
HEALTH TECHNOLOGY ASSESSMENT

VOLUME 22 ISSUE 6 JANUARY 2018
ISSN 1366-5278

Axitinib, cabozantinib, everolimus, nivolumab, sunitinib and best supportive care in previously treated renal cell carcinoma: a systematic review and economic evaluation

Steve J Edwards, Victoria Wakefield, Peter Cain, Charlotta Karner, Kayleigh Kew, Mariana Bacelar, Natalie Masento and Fatima Salih



NHS
National Institute for Health Research

DOI 10.3310/hta22060

Axitinib, cabozantinib, everolimus, nivolumab, sunitinib and best supportive care in previously treated renal cell carcinoma: a systematic review and economic evaluation

Steve J Edwards,* Victoria Wakefield, Peter Cain, Charlotta Karner, Kayleigh Kew, Mariana Bacelar, Natalie Masento and Fatima Salih

BMJ Technology Assessment Group, BMJ, London, UK

*Corresponding author

Declared competing interests of authors: The *British Medical Journal* (BMJ) Technology Assessment Group (TAG) and editorial team of the BMJ work independently of one another. The views and opinions expressed in this report are those of the BMJ-TAG. No competing interests were declared that affect the impartiality of this report.

Published January 2018
DOI: 10.3310/hta22060

This report should be referenced as follows:

Edwards SJ, Wakefield V, Cain P, Karner C, Kew K, Bacelar M, *et al*. Axitinib, cabozantinib, everolimus, nivolumab, sunitinib and best supportive care in previously treated renal cell carcinoma: a systematic review and economic evaluation. *Health Technol Assess* 2018;**22**(6).

Health Technology Assessment is indexed and abstracted in *Index Medicus*/MEDLINE, *Excerpta Medica*/EMBASE, *Science Citation Index Expanded* (SciSearch®) and *Current Contents®*/ Clinical Medicine.

Health Technology Assessment

HTA/HTA TAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 4.236

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the Clarivate Analytics Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: http://www.nets.nihr.ac.uk/programmes/hta

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 16/58/01. The contractual start date was in May 2015. The draft report began editorial review in April 2017 and was accepted for publication in September 2017. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

Abstract

Axitinib, cabozantinib, everolimus, nivolumab, sunitinib and best supportive care in previously treated renal cell carcinoma: a systematic review and economic evaluation

Steve J Edwards,* Victoria Wakefield, Peter Cain, Charlotta Karner, Kayleigh Kew, Mariana Bacelar, Natalie Masento and Fatima Salih

BMJ Technology Assessment Group, BMJ, London, UK

*Corresponding author sedwards@bmj.com

Background: Several therapies have recently been approved for use in the NHS for pretreated advanced or metastatic renal cell carcinoma (amRCC), but there is a lack of comparative evidence to guide decisions between them.

Objective: To evaluate the clinical effectiveness and cost-effectiveness of axitinib (Inlyta®, Pfizer Inc., NY, USA), cabozantinib (Cabometyx®, Ipsen, Slough, UK), everolimus (Afinitor®, Novartis, Basel, Switzerland), nivolumab (Opdivo®, Bristol-Myers Squibb, NY, USA), sunitinib (Sutent®, Pfizer, Inc., NY, USA) and best supportive care (BSC) for people with amRCC who were previously treated with vascular endothelial growth factor (VEGF)-targeted therapy.

Data sources: A systematic review and mixed-treatment comparison (MTC) of randomised controlled trials (RCTs) and non-RCTs. Primary outcomes were overall survival (OS) and progression-free survival (PFS). Secondary outcomes were objective response rates (ORRs), adverse events (AEs) and health-related quality of life (HRQoL). MEDLINE, EMBASE and The Cochrane Library were searched from inception to January and June 2016 for RCTs and non-RCTs, respectively. Two reviewers abstracted data and performed critical appraisals.

Review methods: A fixed-effects MTC was conducted for OS, PFS [hazard ratios (HRs)] and ORR (odds ratios), and all were presented with 95% credible intervals (CrIs). The RCT data formed the primary analyses, with non-RCTs and studies rated as being at a high risk of bias included in sensitivity analyses (SAs). HRQoL and AE data were summarised narratively. A partitioned survival model with health states for pre progression, post progression and death was developed to perform a cost–utility analysis. Survival curves were fitted to the PFS and OS results from the MTC. A systematic review of HRQoL was undertaken to identify sources of health state utility values.

Results: Four RCTs ($n = 2618$) and eight non-RCTs ($n = 1526$) were included. The results show that cabozantinib has longer PFS than everolimus (HR 0.51, 95% CrI 0.41 to 0.63) and both treatments are better than BSC. Both cabozantinib (HR 0.66, 95% CrI 0.53 to 0.82) and nivolumab (HR 0.73, 95% CrI 0.60 to 0.89) have longer OS than everolimus. SAs were consistent with the primary analyses. The economic analysis, using drug list prices, shows that everolimus may be more cost-effective than BSC with an incremental cost-effectiveness ratio (ICER) of £45,000 per quality-adjusted life-year (QALY), as it is likely to be considered an end-of-life treatment. Cabozantinib has an ICER of £126,000 per QALY compared with everolimus and is unlikely to be cost-effective. Nivolumab was dominated by cabozantinib (i.e. more costly and less effective) and axitinib was dominated by everolimus.

Limitations: Treatment comparisons were limited by the small number of RCTs. However, the key limitation of the analysis is the absence of the drug prices paid by the NHS, which was a limitation that could not be avoided owing to the confidentiality of discounts given to the NHS.

Conclusions: The RCT evidence suggests that cabozantinib is likely to be the most effective for PFS and OS, closely followed by nivolumab. All treatments appear to delay disease progression and prolong survival compared with BSC, although the results are heterogeneous. The economic analysis shows that at list price everolimus could be recommended as the other drugs are much more expensive with insufficient incremental benefit. The applicability of these findings to the NHS is somewhat limited because existing confidential patient access schemes could not be used in the analysis. Future work using the discounted prices at which these drugs are provided to the NHS would better inform estimates of their relative cost-effectiveness.

Study registration: This study is registered as PROSPERO CRD42016042384.

Funding: The National Institute for Health Research Health Technology Assessment programme.

Contents

List of tables

List of figures

List of boxes

List of abbreviations

AE	adverse event
AG	assessment group
AIC	Akaike information criterion
amRCC	advanced or metastatic renal cell carcinoma
ASCO	American Society of Clinical Oncology
ASCO-GU	American Society of Clinical Oncology-Genitourinary Cancers Symposium
BIC	Bayesian information criterion
BMJ	*British Medical Journal*
BNF	*British National Formulary*
BSC	best supportive care
CDF	Cancer Drugs Fund
CENTRAL	Cochrane Central Register of Controlled Trials
CI	confidence interval
CiC	commercial in confidence
CRD	Centre for Reviews and Dissemination
CrI	credible interval
CT	computed tomography
CTCAE	common terminology criteria for adverse events
CYP3A4	cytochrome P4SO 3A4
DARE	Database of Abstracts of Reviews and Effects
DSU	Decision Support Unit
ECOG	Eastern Cooperative Oncology Group
EMUC	European Multidisciplinary Meeting on Urological Cancers
EORTC QLQ-C30	European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire
EQ-5D	EuroQol-5 Dimensions
EQ-5D-5L	EuroQol-5 Dimensions, five-level version
EQ-VAS	EuroQoL visual analogue scale
ERG	evidence review group
ESMO	European Society for Medical Oncology
FKSI	Functional Assessment of Cancer Therapy Kidney Symptom Index
FKSI-DRS	Functional Assessment of Cancer Therapy-Kidney Cancer Symptom Index-Disease Related Symptoms
GP	general practitioner
HR	hazard ratio
HRQoL	health-related quality of life
HSUV	health state utility value
HTA	health technology assessment
ICER	incremental cost-effectiveness ratio
IPCW	inverse probability of censoring weighted
IPD	individual patient data
ITT	intention to treat
KM	Kaplan–Meier
MD	mean difference
MeSH	medical subject heading
mRCC	metastatic renal cell carcinoma
MRI	magnetic resonance imaging
MSKCC	Memorial Sloan Kettering Cancer Center
MTA	multiple technology appraisal
MTC	mixed-treatment comparison
mTOR	mammalian target of rapamycin
mTORi	mammalian target of rapamycin inhibitor

NHS EED	NHS Economic Evaluation Database
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NMA	network meta-analysis
NMB	net monetary benefit
OR	odds ratio
ORR	objective response rate
OS	overall survival
OWSA	one-way sensitivity analysis
PAS	patient access scheme
PD	progressed disease
PFS	progression-free survival
PH	proportional hazard
PPES	palmar–plantar erythrodysaesthesia syndrome
PPS	post-progression survival
PRISMA	Preferred Reporting Items for Systematic reviews and Meta-Analyses
PSA	probabilistic sensitivity analysis
PSS	Personal Social Services
QALY	quality-adjusted life-year
RCC	renal cell carcinoma
RCT	randomised controlled trial
RDI	relative dose intensity
RECIST	Response Evaluation Criteria In Solid Tumours
RENCOMP	Renal Comparison
ROBINS-I	Risk Of Bias In Non-randomised Studies – of Interventions
RPLS	reversible posterior leukoencephalopathy syndrome
RPSFTM	rank-preserving structural failure time model
SA	sensitivity analysis
SD	standard deviation
SE	standard error
SEER	Surveillance, Epidemiology and End Results
SSE	sum squared error
STA	single technology appraisal
STC	simulated treatment comparison
TA	technology appraisal
TAG	Technology Assessment Group
TC	terminal care
TEAE	treatment-emergent adverse event
TKI	tyrosine kinase inhibitor
TRAE	treatment-related adverse event
TTD	time to discontinuation
VAS	visual analogue scale
VEGF	vascular endothelial growth factor
WTP	willingness to pay

Note

This monograph is based on the Technology Assessment Report produced for NICE. The full report contained a considerable number of data that were deemed confidential. The full report was used by the Appraisal Committee at NICE in their deliberations. The full report with each piece of confidential data removed and replaced by the statement 'confidential information (or data) removed' is available on the NICE website: www.nice.org.uk.

The present monograph presents as full a version of the report as is possible while retaining readability, but some sections, sentences, tables and figures have been removed. Readers should bear in mind that the discussion, conclusions and implications for practice and research are based on all the data considered in the original full NICE report.

Plain English summary

Renal cell carcinoma (RCC) is the most common type of kidney cancer and more people are diagnosed each year in the UK. Several treatments have recently been developed for patients with RCC that is advanced or has spread to other parts of the body and who have previously had treatment but have worsened. Our review compared the treatments axitinib (Inlyta®, Pfizer Inc., NY, USA), cabozantinib (Cabometyx®, Ipsen, Slough, UK), everolimus (Afinitor®, Novartis, Basel, Switzerland), nivolumab (Opdivo®, Bristol-Myers Squibb, NY, USA), sunitinib (Sutent®, Pfizer, Inc., NY, USA) and best supportive care (BSC) to help NHS services choose the most effective option.

The review found that cabozantinib is probably the best treatment to delay tumour growth and prolong life, followed by nivolumab. All of the treatments delayed tumour growth compared with BSC; however, there are uncertainties, owing to the way in which studies have been conducted. All of the treatments cause serious side effects and so it is important that the possible benefits and harms are discussed fully with a cancer specialist before a patient starts treatment. Standard reporting of the most important outcomes for people with RCC, particularly the response to treatment and quality of life, would improve our knowledge of how these treatments compare with each other.

The publicly available prices for these drugs are very high and would require a significant improvement in survival and/or quality of life for them to be considered as cost-effective. The results show that only everolimus had a large enough improvement in survival in comparison with BSC to be cost-effective. However, all of the drugs have commercially confidential discounts for the NHS and so the results of this review are unlikely to be accurate.

Scientific summary

Background

Treatments for advanced or metastatic renal cell carcinoma (amRCC) aim to prevent or slow further spread by targeting pathways that support tumour growth. Patients who do not tolerate first-line treatments or whose disease has progressed may require subsequent therapy. There is a lack of evidence about the relative clinical effectiveness and cost-effectiveness to guide decisions between treatments, several of which have recently been approved for use in the NHS for second-line, and above, treatments.

Objectives

The objectives of this systematic review are to:

- evaluate the clinical effectiveness and cost-effectiveness of axitinib (Inlyta®, Pfizer Inc., NY, USA), best supportive care (BSC), cabozantinib (Cabometyx®, Ipsen, Slough, UK), everolimus (Afinitor®, Novartis, Basel, Switzerland), nivolumab (Opdivo®, Bristol-Myers Squibb, NY, USA), and sunitinib (Sutent®, Pfizer, Inc., NY, USA) for treated amRCC in line with their respective marketing authorisations
- identify key areas for further primary and secondary research.

The review focuses on patients who have received prior vascular endothelial growth factor (VEGF)-targeted therapy and not prior cytokines to reflect treatment sequences in UK clinical practice.

Methods

A systematic review was undertaken to compare the clinical effectiveness of treatments using mixed-treatment comparison (MTC). Randomised controlled trial (RCT) data were preferred but did not link all treatments in the network and so non-RCTs were sought to link in axitinib and sunitinib for the primary outcomes. Studies comparing treatments of interest with sorafenib (Nexavar®, Bayer, Leverkusen, Germany) were also included to connect the network. Eligible studies compared two or more treatments of interest for people with amRCC previously treated with VEGF-targeted therapy. Placebo has been used as a surrogate for BSC. Primary outcomes were overall survival (OS) and progression-free survival (PFS). Secondary outcomes were objective response rate (ORR), adverse events of treatment and health-related quality of life (HRQoL).

The databases MEDLINE, EMBASE and The Cochrane Library were searched from inception to January and June 2016 for RCTs and non-RCTs, respectively. Additional searches were conducted of reference lists of included studies and systematic reviews, conference abstracts and trial registries for ongoing studies. Two or more reviewers sifted the searches, reviewed full papers, abstracted study data and performed critical appraisals.

Fixed-effects MTCs using Bayesian Markov chain Monte Carlo simulation were conducted for OS, PFS and ORR. Primary analyses were limited to RCT data and hazard ratio (HR) (OS and PFS) or odds ratio (OR) (the ORR), with associated 95% credible intervals (CrIs), were used as summary statistics. Non-RCTs and studies rated as being at a high risk of bias were included in sensitivity analyses (SAs). Subgroup analyses to explore the effect of prior therapies and baseline prognostic scores were also carried out for OS and PFS. HRQoL and adverse event (AE) data were summarised narratively owing to inconsistencies in data reporting.

An additional systematic review was undertaken to search for published cost-effectiveness analyses, costing studies and quality-of-life studies in patients with amRCC. A review of National Institute for Health and Care Excellence (NICE) technology appraisals (TAs) of second-line treatments for amRCC was also undertaken.

A cost–utility analysis comparing axitinib, BSC, cabozantinib, everolimus and nivolumab was performed by developing a partitioned survival model in Microsoft Excel® (Microsoft Corporation, Redmond, WA, USA). The model consisted of three health states: pre progression, post progression and death. Outcomes were calculated at each 2-weekly cycle up to a time horizon of 30 years. The perspective was reflective of the NHS in England.

To estimate the expected proportion of patients in each health state at each cycle, parametric survival curves were fitted to digitised Kaplan–Meier data taken from published plots for PFS and OS from the CheckMate 025 trial (Motzer RJ, Escudier B, McDermott DF, George S, Hammers HJ, Srinivas S, *et al.* Nivolumab versus everolimus in advanced renal-cell carcinoma. *N Engl J Med* 2015;**373**:1803–13). This enabled the proportions to be calculated for the nivolumab and everolimus groups, while HRs from the MTC were applied to estimate the proportions for each of the remaining treatments.

A range of scenario analyses were performed, as were probabilistic and deterministic SAs.

Results

Twelve studies ($n = 4144$) met the inclusion criteria: four RCTs (one double-blind RCT and three open-label RCTs) and eight non-RCTs (six retrospective cohort studies and two crossover RCTs in which only second-phase data were relevant). Populations were predominantly male and white, and the mean age was generally between 60 and 70 years. When reported, most patients had stage 3 or 4 clear-cell renal cell carcinoma (RCC) and reasonably good baseline performance status.

The primary PFS analysis, based on two RCTs (RECORD-1 and METEOR), included cabozantinib, everolimus and BSC and showed statistically significant benefits for cabozantinib and everolimus compared with BSC (HR 0.17, 95% CrI 0.12 to 0.24; and HR 0.33, 95% CrI 0.25 to 0.43, respectively), and for cabozantinib compared with everolimus (HR 0.51, 95% CrI 0.41 to 0.63).

A SA for PFS connected axitinib and sunitinib by including five non-RCTs and a third RCT; this analysis showed statistically significant benefits of all active treatments compared with BSC (everolimus HR 0.33, 95% CrI 0.25 to 0.43; cabozantinib HR 0.17, 95% CrI 0.12 to 0.24; axitinib HR 0.31, 95% CrI 0.21 to 0.44; and sunitinib HR 0.27, 95% CrI 0.17 to 0.40). Cabozantinib showed a statistically significant benefit compared with all other treatments: everolimus (HR 0.51, 95% CrI 0.41 to 0.63), sunitinib (HR 0.63, 95% CrI 0.44 to 0.95), axitinib (HR 0.54, 95% CrI 0.40 to 0.76) and BSC (HR 0.17, 95% CrI 0.12 to 0.24). None of the differences in PFS between sunitinib, everolimus and axitinib was statistically significant. Cabozantinib was found to have a 99% probability of being the most effective treatment for improving PFS. Data were not available to provide a robust estimate of PFS for nivolumab compared with other treatments.

The primary OS analysis, based solely on RCT data, included cabozantinib, everolimus, nivolumab and BSC, and did not show statistically significant benefits for any treatment compared with BSC. This is likely to be due to uncertainty in the efficacy of BSC caused by RECORD-1 (Cella D, Michaelson MD, Bushmakin AG, Cappelleri JC, Charbonneau C, Kim ST, *et al.* Health-related quality of life in patients with metastatic renal cell carcinoma treated with sunitinib vs interferon-alpha in a phase III trial: final results and geographical analysis. *Br J Cancer* 2010;**102**:658–64) requiring crossover adjustment for confounding. All mean estimates were in favour of the active treatments. Cabozantinib and nivolumab led to longer OS compared with everolimus (HR 0.66, 95% CrI 0.53 to 0.82; and HR 0.73, 95% CrI 0.60 to 0.89, respectively); however, the difference between nivolumab and cabozantinib was not statistically significant (HR 1.12, 95% CrI 0.82 to 1.49). Cabozantinib was associated with the highest probability of being the best treatment for this outcome (72%).

The results of the SA for OS, including data to compare treatments with axitinib, were in keeping with those of the primary analysis. Everolimus, cabozantinib and nivolumab showed longer OS compared with axitinib (HR 0.74, 95% CrI 0.56 to 0.99; HR 0.48, 95% CrI 0.34 to 0.71; and HR 0.54, 95% CrI 0.38 to 0.77, respectively). Data were not available to provide an OS estimate for sunitinib compared with the other treatments and there was statistically significant inconsistency in the network for this SA.

The primary ORR analysis, based on three RCTs including cabozantinib, everolimus, nivolumab and BSC, showed statistically significant benefits of all treatments compared with BSC. Cabozantinib and nivolumab resulted in statistically significant improvements in ORR compared with everolimus (OR 6.67, 95% CrI 3.28 to 12.78; and OR 6.18, 95% CrI 3.75 to 9.84, respectively). The difference between nivolumab and cabozantinib was not statistically significant for ORR (OR 1.05, 95% CrI 0.41 to 2.18). CheckMate 025 (nivolumab vs. everolimus) was rated as being at a high risk of bias owing to the absence of blinding of outcome assessors for response and METEOR (cabozantinib vs. everolimus) [Choueiri TK, Escudier B, Powles T, Tannir NM, Mainwaring PN, Rini BI, *et al.* Cabozantinib versus everolimus in advanced renal cell carcinoma (METEOR): final results from a randomised, open-label, phase 3 trial. *Lancet Oncol* 2016;**17**:917–27] was rated as being at an unclear risk of bias for missing data, but the impact of these potential biases on the overall direction of treatment effects is unknown.

Treatments could not be compared using MTC for HRQoL as different measures and tools were used for assessments. HRQoL scores were similar between axitinib and sorafenib in AXIS (Motzer RJ, Escudier B, Tomczak P, Hutson TE, Michaelson MD, Negrier S, *et al.* Axitinib versus sorafenib as second-line treatment for advanced renal cell carcinoma: overall survival analysis and updated results from a randomised phase 3 trial. *Lancet Oncol* 2013;**14**:552–62) and results favoured nivolumab over everolimus in CheckMate 025. Results in RECORD-1 favoured BSC over everolimus, although this effect was only apparent if models were used to account for data not missing at random. METEOR results were similar for everolimus and cabozantinib. The way that AEs were defined and categorised differed across studies and, therefore, no MTC could be conducted and the narrative synthesis was limited to RCT data. The rate of grade 3/4 AEs was higher with everolimus (36.5%) than nivolumab (18.7%), and higher with cabozantinib (71.0%) than everolimus (59.9%).

In the assessment of cost-effectiveness, the base-case analysis showed that nivolumab was the most expensive treatment overall, followed closely by cabozantinib, at £107,000 and £106,000, respectively. Axitinib incurred a significantly lower cost of £46,000, while everolimus totalled £36,000. BSC had a mean overall cost of £11,000.

The most effective treatment in the base-case analysis was cabozantinib, which accrued a mean of 1.87 quality-adjusted life-years (QALYs) over the time horizon of the model. This was noticeably more effective than nivolumab, which accrued 1.60 QALYs, which was markedly better than everolimus and axitinib, which each accrued 1.31 QALYs. BSC accrued just 0.75 QALYs over the time horizon. These values were largely driven by differences in OS, which led cabozantinib to have an expected mean survival of 3.18 years compared with just 2.53 for nivolumab. Everolimus and axitinib both resulted in a mean of 2.21 life-years due to the assumption that axitinib was as effective as everolimus, and BSC had an associated survival of 1.25 years.

These results mean that everolimus dominated axitinib as it accrued the same number of QALYs but incurred a lower cost, while nivolumab was dominated by cabozantinib, which accrued more QALYs at a slightly lower cost. The incremental analysis then simplifies to a comparison between cabozantinib, everolimus and BSC, resulting in an incremental cost-effectiveness ratio (ICER) of £45,000 per QALY for everolimus compared with BSC, and £126,000 per QALY for cabozantinib compared with everolimus.

A probabilistic SA was performed with 10,000 samples, resulting in similar results of £45,000 per QALY for everolimus compared with BSC, and £123,000 per QALY for cabozantinib compared with everolimus. Deterministic SAs showed that the most sensitive parameters were the OS HR and relative dose intensity (RDI) for the active treatments. When the upper values were used for the OS HRs relative to everolimus, the

ranking changed so that everolimus became optimal at a threshold of £30,000 per QALY. The same was also true for the lower bound of the RDI for everolimus but, when the upper bound was used, BSC remained optimal and axitinib became preferable to everolimus. Axitinib also became preferable to everolimus when the upper RDI value was used for axitinib. The ranking of nivolumab and cabozantinib changed when the lower RDI value for nivolumab or the upper RDI value of cabozantinib was used, resulting in nivolumab being preferable to cabozantinib at a threshold of £30,000 per QALY.

A range of scenario analyses were performed including varying the distributions applied for OS and using the HRs derived from an extended network in the MTC to include CheckMate 025 as well as the identified observational evidence. Axitinib and nivolumab were dominated in all scenarios as with the base-case analysis, and the ICERs for everolimus compared with BSC ranged from £38,000 to £46,000 per QALY, in comparison with the base-case ICER of £45,000 per QALY. For cabozantinib compared with everolimus, the ICERs ranged from £102,000 per QALY to £248,000 per QALY, in comparison with the base-case ICER of £126,000 per QALY.

Discussion

This review was conducted according to robust methods that were prespecified in a prospectively registered protocol. The primary analyses bring together high-quality evidence from RCTs for the most pertinent outcomes in this population, using MTC when possible to estimate relative treatment effects in the absence of head-to-head evidence. The inclusion criteria were widened to incorporate comparative observational evidence in SAs to substantiate the primary results and to provide estimates for all treatments of interest.

Treatment comparisons were limited by a small number of RCTs. The proportional hazards assumption did not hold for PFS in CheckMate 025, which prevented the inclusion of nivolumab; randomised evidence for axitinib was limited to a subgroup analysis of AXIS that did not connect to the other RCTs in the network; and imprecision surrounding BSC (informed by RECORD-1) led to counterintuitive results in the OS analysis. SAs incorporating non-randomised evidence provided relative effects for more treatments, but introduced inconsistency and probably bias.

Planned subgroup analyses for prior therapies and baseline prognostic score could not provide results for all treatments and there were too few studies informing the MTC to support additional analyses to explore whether or not observed inconsistencies [e.g. everolimus AE rates in METEOR (59.9%) and RECORD-1 (36.5%)] could be explained by design or between-group baseline differences.

The main limitation of the review is that the costs are based on the list prices of the drugs. There are patient access schemes (PASs) in place to provide these drugs on the NHS with a reduced price or pricing strategy. The details of these PASs are confidential and so could not be incorporated in the analysis. This limits the applicability of the results, which may not reflect current practice in the UK. A strength of the analysis is the range of models tested to fit survival models, which included flexible spline models that proved to have a very good fit to the PFS data in CheckMate 025. The analysis explored a range of scenarios to test different assumptions on the base-case results.

A range of SAs was performed including a probabilistic analysis with a large number of samples. The robustness of the results has therefore been thoroughly tested and the model was found to only be sensitive to a few key parameters: the relative OS of treatments and the RDI for each treatment. These results are not surprising as OS is an influential driver on the total QALYs by definition and the RDI has an impact on the treatment acquisition costs, which make up the majority of overall treatment costs.

Conclusions

The current evidence base to inform decisions between axitinib, cabozantinib, everolimus, nivolumab, BSC and sunitinib for previously treated amRCC is limited by the number of studies providing comparative clinical effectiveness data, and by the quality of study reporting. Analyses of PFS and OS suggest that cabozantinib is likely to be the most effective treatment, closely followed by nivolumab, and with little difference between axitinib, everolimus and sunitinib. All treatments considered in this review appear to delay disease progression and prolong survival more than providing BSC. Cabozantinib is not yet available for use in the NHS in England, although it is currently undergoing appraisal by NICE.

High-quality RCT data comparing all the available RCC treatment options are required to enable more robust estimates of efficacy, including RCTs comparing newer RCC therapies with more established treatments. Further PFS data from a RCT are also required for nivolumab to enable its inclusion in a MTC and more standardised reporting of response rates, HRQoL (e.g. EuroQol-5 Dimensions) and AEs in RCTs would facilitate direct comparisons of the RCC treatments.

The economic analysis showed that the majority of current treatments for second-line RCC are very expensive and unlikely to be cost-effective at list price. The exception to this is everolimus, which may be cost-effective at the NICE threshold of £50,000 per QALY granted to treatments that qualify as an end-of-life treatment. All drugs assessed in this analysis have confidential PASs that provide them to the NHS at a discounted price. The economic results may therefore not fully reflect the current NHS setting and should be considered with caution.

Study registration

This study is registered as PROSPERO CRD42016042384.

Funding

Funding for this study was provided by the Health Technology Assessment programme of the National Institute for Health Research.

Chapter 1 Background

Description of the health problem

Kidney cancer is the seventh most common cancer in the UK, accounting for 3% of all new cancer cases.[1] From 2016, incidence is projected to increase by at least 25% by 2035, making it one of the fastest accelerating cancers in the UK.[2] Kidney cancers are more common in men and older people, but incidence is rising most sharply in women.[3] At least four out of every five kidney cancers in the UK are renal cell carcinomas (RCCs), which originate in cells lining the tubules filtering waste from the blood to the urine.[4] The 10-year survival for kidney cancer of any type in the UK is 50%, although this varies in particular by stage of cancer at diagnosis.[1]

Three important risk factors for RCC are smoking, obesity and germline mutations, which contribute to about 42% of all kidney cancer cases in the UK.[1,5] Hypertension and advanced kidney disease also increase the risk of RCC and are associated with a worse prognosis.[5] Specific dietary habits, occupational exposure to carcinogens (e.g. asbestos), certain medical conditions and medications, and a sedentary lifestyle have also been implicated.

Renal cell carcinomas, like most cancers, are usually described by numerical stages from I to IV, which helps determine appropriate treatment. Stage I and II tumours are both located completely inside the kidney, the latter being > 7 cm across. Surgery to remove the tumour is the main treatment for RCC at these stages. Stage III tumours may have spread to a major vein or into tissue around the kidney and may involve one nearby lymph node. If the RCC has spread further into the surrounding tissue and involves more than one lymph node or has spread to other parts of the body, the cancer is termed metastatic (stage IV).[6] The main focus of drug therapies for stage III and IV RCC is to prevent or slow further growth, but surgery may also be appropriate to remove the primary and secondary tumours. Five-year survival is > 80% for people with stage I disease and < 10% for people with stage IV disease.[7]

In addition to tumour stage, RCC can be classified by cell histology. Histological variants have distinctive cell appearance under a microscope and vary by the stage they are likely to be diagnosed, their incidence pattern across age and sex, and their prognosis.[4,8] Clear-cell RCC is by far the most common, accounting for around 80% of RCC cases. Other variants fall under the umbrella term of non-clear cell, but vary significantly. Within these, papillary RCC accounts for around 10% of cases, chromophobe RCC about 5%, and collecting duct carcinoma around 1%.[4,8] Papillary and chromophobe RCC tend to have a more favourable prognosis than clear-cell RCC and collecting duct tends to have a less favourable prognosis than clear-cell RCC.[8] Several rare variants have been identified and around 5% of cases cannot be classified. Prognosis is worse if the tumour becomes sarcomatoid, which can occur in any of the variants.[9,10]

In the UK, the most recent data show 7800 new male cases per year and an age-standardised rate of approximately 14 per 100,000 people (2008 to 2010),[3] compared with 4700 new female cases with a rate of approximately eight per 100,000 people. Incidence has increased significantly in recent decades and continues to rise,[2] which is thought to be explained by an ageing population, increases in obesity and more widespread use of cross-sectional imaging, which results in the detection of asymptomatic incidental cancers.[3] Rising incidence, which is more pronounced in women than in men, has not been mirrored by increases in kidney cancer mortality rates.[3] Incidence rises sharply from around the age of 45–49 years and peaks between 85 and 89 years; about 50% of all kidney cancer cases in the UK are diagnosed in those aged ≥ 70 years.[3] Kidney cancer survival in England is highest for those who were diagnosed before the age of 50 years (2009–13).[7] No recent prevalence data for RCC in England were identified.

Early-stage kidney cancer may not cause any symptoms, meaning that stage I and II tumours are regularly picked up during routine medical investigations (> 50%), and 25–30% of cases present at stage IV.[11,12] If symptoms are present, these may include blood in the urine (which may not be identified until testing), a lump or mass in the kidney area, and localised flank pain. If RCC is suspected, diagnosis is usually made by ultrasonography, computed tomography (CT) or magnetic resonance imaging (MRI), but sometimes a biopsy is required to confirm.[13] Less common and non-specific symptoms may include fatigue, loss of appetite or weight, recurrent fevers, persistent side pain, high blood pressure and anaemia. Symptoms of metastases include bone pain or lung nodules, hypercalcaemia, unexplained fever, erythrocytosis and wasting syndromes.[11]

Rating scales can be used for RCC to assess a range of factors associated with disease status and predicted survival [e.g. Memorial Sloan Kettering Cancer Center (MSKCC) or Heng criteria].[14,15] These include a measure of the extent to which a person can engage in usual daily activities [e.g. the Karnofsky or Eastern Cooperative Oncology Group (ECOG) scales],[16,17] timing of diagnosis and treatment, and blood markers (e.g. haemoglobin, calcium, platelets and neutrophils). Scoring across these domains allows RCC risk to be categorised from favourable to poor on a scale of worsening predicted survival.

Renal cell carcinoma has a serious effect on patients' physical, social and psychological well-being, particularly when it is advanced or metastatic.[12] Symptoms and treatment toxicity are a significant physical burden for patients and their caregivers as RCC progresses, compounded by the social and psychological effects of living with advanced cancer. Once metastatic, surgery is rarely an option, resistance to targeted therapies is common and median survival is < 1 year.[12] As the incidence of RCC rises in the UK, owing to an ageing population and the rising rate of obesity, the burden on the NHS is set to increase significantly.[2,18]

Annual NHS costs for cancer services are > £5B and wider societal costs, including the economic impact of premature death and loss of productivity, have been estimated at £18.3B.[19] There are no UK cost-of-illness data to estimate how much RCC contributes to this economic burden, but the cost of emerging targeted therapies for first- and second-line treatment of advanced or metastatic renal cell carcinoma (amRCC), along with increasing incidence relative to other cancers, means its share is likely to be increasing. Considering all causes of UK death and disability, renal cancers account for around 1 in every 133 years of life lost, 1 in every 169 deaths, and 1 in every 227 disability-adjusted life-years.[20]

Current service provision

Treatments for stage I and II RCC aim to remove the tumour (full or partial nephrectomy) or shrink it, either by ablation (radiofrequency or cryotherapy ablation) or by cutting off its blood supply (embolisation).[21] In addition to treatment of the tumour, the National Institute for Health and Care Excellence (NICE) provides guidance on improving supportive and palliative care within the care pathway for all cancers.[22] Depending on the individual's needs, this may include psychological and social support, rehabilitation, complementary therapy services and support for families and carers.

In advanced and metastatic RCC (stage III and IV), the aim of treatment is to slow the growth or spread of the cancer, usually with vascular endothelial growth factor (VEGF)-targeted therapies. Since the emergence of these therapies, cytokines (interleukin 2 or interferon) are no longer commonly used for advanced RCC, primarily owing to their association with severe adverse events (AEs) (e.g. myocardial infarction, intestinal bleeding and kidney damage). Current first-line therapies recommended by NICE for initial treatment of stage III or IV RCC are the tyrosine kinase inhibitors (TKIs) pazopanib (Votrient®, Novartis, Camberley, UK) and sunitinib (Sutent®, Pfizer Inc., NY, USA).[23,24] These are indicated for patients who have not received prior cytokine therapy and have an ECOG performance status of 0 or 1.

If a patient does not respond to, is intolerant of or progresses on the first TKI (or in some cases on a prior cytokine), NICE currently recommends axitinib (Inlyta®, Pfizer Inc., NY, USA),[25] everolimus (Afinitor®,

Novartis, Basel, Switzerland)[26] or nivolumab (Opdivo®, Bristol-Myers Squibb, NY, USA) as second-line treatment.[27] Cabozantinib (Cabometyx®, Ipsen, Slough, UK) has recently received marketing authorisation and is currently undergoing appraisal by NICE for the same indication.[28] The marketing authorisation for sunitinib is not limited to untreated patients, but NICE has not recommended it for second-line use for people with amRCC.[29] Axitinib, cabozantinib and sunitinib are oral TKIs, everolimus is an oral mammalian target of rapamycin inhibitor (mTORi), and nivolumab is a human monoclonal antibody given intravenously. Sorafenib (Nexavar®, Bayer, Leverkusen, Germany) has a UK marketing authorisation for the treatment of people with advanced RCC who have received (or are unsuitable for) interferon-alpha or interleukin 2-based therapy, but it is not recommended by NICE for first- or second-line treatment.[29]

Regional variations in the percentage of patients receiving second-line treatment were observed in the UK RECCORD registry of RCC patients who started treatment between 2009 and 2012.[30] Overall, 15.8% of patients in the registry received a second-line therapy, although the proportion was substantially higher in England than in Wales or Scotland (19.5%, 7.5% and 8.5%, respectively). At the time, everolimus (53.1%), sunitinib (14.8%) and pazopanib (9.9%) were the most commonly used second-line treatments,[30] but these data pre-date the approval of axitinib and nivolumab for pre-treated RCC,[25,31] and the approval of everolimus for routine use in the NHS rather than through the Cancer Drugs Fund (CDF).[26] There is currently no NICE clinical guideline for the treatment of RCC and although evidence from individual technology appraisals (TAs) is summarised in a NICE pathway,[32] this does not provide guidance for decisions between the available treatments.

Description of the technologies under assessment

This systematic review will consider evidence of the clinical effectiveness and cost-effectiveness of targeted therapies for adults with amRCC who have received previous VEGF-targeted therapy. The therapies being assessed in this review are axitinib, cabozantinib, everolimus, nivolumab and sunitinib (*Table 1*). Axitinib, cabozantinib and sunitinib are oral TKIs, a group of targeted cancer drugs that suppress cancer progression by inhibiting growth proteins (tyrosine kinases) of tumour cells and their associated blood supply. Everolimus is an oral mTORi, a drug class that also target cell division and tumour blood supply but via the inhibition of a different growth regulator protein called mammalian target of rapamycin (mTOR). Nivolumab is an intravenous immunotherapy. It is a human monoclonal antibody, which induces a targeted immune response to cancer cells by blocking an immune checkpoint protein receptor called programmed cell death protein 1.

TABLE 1 Summary of technologies under review

Generic	Brand	Company	Class	Route	Available as	Standard regimen
Axitinib	Inlyta®	Pfizer Inc. (NY, USA)	TKI	Oral	1-, 3-, 5- and 7-mg tablets	5 mg b.i.d.
Cabozantinib	Cabometyx®	Ipsen (Paris, France)	TKI	Oral	20-, 40- and 60-mg tablets	60 mg q.i.d.
Everolimus	Afinitor®	Novartis (Basel, Switzerland)	mTORi	Oral	2.5-, 5-, 10-mg tablets	10 mg q.i.d.
Nivolumab	Opdivo®	Bristol-Myers Squibb (NY, USA)	mAb	i.v.	10-mg/ml concentrate for solution for infusion	3 mg/kg/2 weeks
Sunitinib	Sutent®	Pfizer Inc. (NY, USA)	TKI	Oral	12.5-, 25-, 37.5- and 50-mg tablets	50 mg q.i.d., 4 weeks on, 2 weeks off cycle

b.i.d., twice daily; i.v., intravenous; mAb, monoclonal antibody; q.i.d., once daily.

Hypersensitivity and toxicity have been observed for all the medicines being assessed meaning dose adjustment or discontinuation may be necessary and regular monitoring is required alongside routine cancer care.[25] Patients may also require additional treatment to prevent or manage treatment-related adverse reactions. There is a large degree of overlap in the most commonly reported adverse reactions (e.g. fatigue, nausea, diarrhoea, stomatitis and rash) but the drugs differ in their contraindications and rarer, more serious adverse reactions outlined below.

Axitinib is a TKI administered orally as a 5-mg tablet twice daily. It has a marketing authorisation in the UK for the treatment of adults with advanced RCC after failure of previous treatment with sunitinib or a cytokine.[33] Patients taking axitinib require regular follow-up to monitor for AEs including thyroid dysfunction, cardiac events, gastrointestinal perforation and fistula formation, proteinuria and liver-related reactions. Hypertension is commonly reported and should be closely monitored, particularly during the first month of treatment. Axitinib should be used with caution in patients with a history of arterial and venous thrombolytic events, and should not be used in those with untreated brain metastases or recent gastrointestinal bleeding. Axitinib has potential wound healing implications that would require caution or temporary cessation if surgery is indicated.

Cabozantinib is a TKI also administered orally. The standard dose is 60 mg daily but dose adjustments or temporary interruption may be required in the event of unacceptable toxicity.[34] As such, close evaluation is recommended for the first 8 weeks when events are most likely to occur. Cabozantinib should be used with caution in patients with mild to moderate renal impairment (not recommended for severe) and those with a history of QT interval prolongation. Careful evaluation is also required for patients who have recently received radiotherapy or surgery, or have gastrointestinal tumour infiltration or inflammatory bowel disease, as there is an increased risk of serious gastrointestinal perforations, fistulas and intra-abdominal abscesses.[34] Other serious AEs that required close monitoring, and on some occasions discontinuation, during cabozantinib treatment included haemorrhage, pneumonia, mucosal inflammation, palmar–plantar erythrodysaesthesia syndrome (PPES), reversible posterior leukoencephalopathy syndrome (RPLS), wound complications, hypertension, proteinuria and venous or arterial thromboemolytic events. Cabozantinib interacts with cytochrome P4SO 3A4 (CYP3A4) inhibitors and inducers, gastric pH-modifying agents, P-glycoprotein substrates, multidrug resistance-associated protein 2 (MRP2) inhibitors, and bile salt-sequestering agents.[34]

Everolimus is an mTORi administered orally, usually as a 10-mg tablet once daily.[33] It has a marketing authorisation in the UK for the treatment of people with advanced RCC after treatment with VEGF-targeted therapy. Common adverse reactions observed during clinical trials were stomatitis, rash, fatigue, diarrhoea, infections, nausea, decreased appetite, anaemia, dysgeusia, pneumonitis, peripheral oedema, hyperglycaemia, asthenia, pruritus, weight loss, hypercholesterolaemia, epistaxis, cough and headache. People taking everolimus require close monitoring for potential severe, and sometimes fatal, adverse reactions, including non-infectious pneumonitis, immunosuppression, renal failure and hypersensitivity reactions, including anaphylaxis, dyspnoea, chest pain and angiooedema.[33] Coadministration with CYP3A4 inducers or multidrug efflux pump P-glycoprotein should be avoided, and those taking angiotensin-converting enzyme inhibitors may be particularly at risk of angiooedema. Caution is recommended for patients with mild to moderate hepatic impairment and in the pre-surgical period owing to wound healing complications with this class of medicine.[33]

Nivolumab is a human monoclonal antibody administered by intravenous infusion at a dose of 3 mg/kg over 60 minutes every 2 weeks,[34] which involves staff and infrastructure costs not required for the oral treatments. It has a UK marketing authorisation for adults with advanced RCC after prior therapy. Dose escalation or reduction is not recommended, but delay or discontinuation may be required in the event of severe immune-related adverse reactions such as pneumonitis, hepatitis, colitis, nephritis and endocrinopathies. Systemic corticosteroids and other immunosuppressants should be avoided before starting nivolumab, owing to their potential interference with nivolumab pharmacodynamic activity, but may be required to treat immune-related reactions. Common and very common AEs associated with nivolumab are fatigue, rash, pruritus, diarrhoea, nausea, respiratory infections and reactions, hypertension, dry eye, peripheral neuropathy, headache, dizziness, decreased appetite, and neutropenia.[34] Patients receiving nivolumab require regular monitoring for at least 5 months after the last dose as adverse

reactions may occur at any time during or after discontinuation. It can be administered as combination therapy with ipilimumab (Yervoy®, Bristol-Myers Squibb, Uxbridge, UK) for some indications, but only nivolumab monotherapy will be considered in this review.

Sunitinib is a TKI administered orally as a 50-mg oral tablet once daily for 4 weeks followed by 2 weeks off, and repeated in a 6-week cycle.[35] Sunitinib has a UK marketing authorisation for the treatment of amRCC but it is not recommended by NICE for second-line treatment.[29] Skin and hair discolouration, bleeding and haemorrhage events, hypertension, anaemia and gastrointestinal reactions are observed commonly and require close monitoring and regular complete blood counts, particularly in those with associated medical histories. Routine monitoring of thyroid function, urinalysis and glucose levels are also recommended.[35] Cases of renal impairment, thromboembolic and pulmonary events, fistula formation, impaired wound healing, dysgeusia, cardiac events, QT interval prolongation, seizures and RPLS, and serious infection have been reported; caution should be exercised with sunitinib for patients with a history, or at higher risk of, these events. Concomitant use of sunitinib with potent CYP3A4 inhibitors or intravenous bisphosphonates should be avoided because of increased plasma levels and the risk of osteonecrosis of the jaw, respectively.[35]

Chapter 2 Definition of the decision problem

The treatment options available to clinicians and their patients with amRCC at second-line treatment and beyond have changed substantially over the last few years. In particular, marketing authorisation has been granted for new therapies such as nivolumab and cabozantinib; everolimus, which was previously only available through the CDF, is now recommended by NICE for routine use.[26] The treatment pathway for advanced and metastatic RCC has changed owing to more treatment options becoming available for patients. However, the evidence for the use of newer treatments is generally limited to the trials used to gain the regulatory approval, with no direct head-to-head randomised controlled trial (RCT) data to evaluate how they compare with other new treatments or other older established treatments. These changes have highlighted the need for a UK-based review summarising the clinical effectiveness and cost-effectiveness of the currently available treatment options, to help inform clinical practice and decision-making.

The original protocol for this review was designed in liaison with NICE as the review was planned as a multiple technology appraisal (MTA).[36] However, the MTA was suspended shortly after the completion of a final protocol owing to changes to the technologies that were due to be appraised within the NICE single technology appraisal (STA) programme.[37] The comparators in the NICE MTA were axitinib, sorafenib and sunitinib in previously treated RCC. The comparators included in the protocol for this review, which has been commissioned by National Institute for Health Research (NIHR), were axitinib, best supportive care (BSC), everolimus, nivolumab, sorafenib and sunitinib. It should be noted that there have been several changes from the original protocol to reflect the changes in current practice and these will be discussed in detail below.

Decision problem

The final inclusion criteria for the review are detailed in *Table 2*. In summary, the review considers comparative effectiveness data for axitinib, BSC, cabozantinib, everolimus, nivolumab and sunitinib in people who had received prior VEGF-targeted therapy for amRCC. It was planned that the review would only consider RCT evidence, but this was expanded to include comparative observational studies to link all treatments of interest.

Population

The population of interest in this review was people who had received at least one prior VEGF-targeted therapy for amRCC. The final protocol included a second population: people who had received at least one prior cytokine therapy for amRCC. However, following feedback from clinical experts in the UK, it was deemed that this population was no longer of relevance. It should be noted that the decision to remove the prior cytokine population from the review was made following the primary searches. As a result, the search strategies and initial abstract appraisal were broader than necessary, although all studies relating only to patients who had received prior cytokines were subsequently excluded from the review and accounted for in the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) diagram presented in *Chapter 3, Results*.

Interventions

The interventions of interest in this review are axitinib, cabozantinib, everolimus, nivolumab and sunitinib for treated amRCC in line with their respective marketing authorisations. Cabozantinib was added after completion of the protocol because it received UK marketing authorisation for use in RCC and was due to be appraised by NICE (TA463).[28,38] A NICE appraisal of lenvatinib (Kisplyx®, Eisai Co., Ltd, Tokyo, Japan) with everolimus for this indication was in process during the writing of this report; however, the associated UK marketing authorisation had not been granted at this time so the treatment was not included.

TABLE 2 Inclusion criteria

PICO criteria	Inclusion criteria as listed in protocol	Final review inclusion criteria	Summary of changes (if any)
Study design	RCTs (comparative non-RCTs will be considered when RCT evidence is insufficient to inform decision problem)	RCTs (comparative non-RCTs will be considered when RCT evidence is insufficient to inform decision problem)	N/A
Population	Patients with previously treated amRCC	Patients with previously treated amRCC	N/A
Interventions	For patients who have received previous cytokine therapy (aldesleukin or interferon alfa): • axitinib • sorafenib • sunitinib • BSC For people who have received previous VEGF-targeted therapy: • axitinib • everolimus • nivolumab • sunitinib	For people who have received previous VEGF-targeted therapy: • axitinib • cabozantinib • everolimus • nivolumab • sunitinib	Previous cytokine therapy population removed as it is no longer used in first-line treatment of advanced RCC in the UK. Cabozantinib added as it has now received marketing authorisation for use in the UK and is currently undergoing appraisal by NICE (TA463)[28]
Comparators	• The interventions listed above compared with each other • BSC	• The interventions listed above compared with each other • BSC	N/A
Outcome	• Overall survival • Progression-free survival • Response rates • Adverse events of treatment • HRQoL	• Overall survival • Progression-free survival • Response rates • Adverse events of treatment • HRQoL	N/A

HRQoL, health-related quality of life; N/A, not applicable.

In addition to axitinib and sunitinib, sorafenib was an intervention of interest in only the subgroup of people who had received prior cytokines. This is because sorafenib is only licensed for use in the UK in patients with advanced RCC who have failed prior interferon-alpha or interleukin 2-based therapy, or are considered unsuitable for such therapy.[39] Hence, following the removal of the population of people who had received only prior cytokines from the review question, sorafenib was no longer an intervention of interest.

Sunitinib was listed as an intervention of interest in the final protocol and has been included in the report although it should be noted that it is only recommended by NICE for use at first line in amRCC (TA169).[24] Sunitinib is not recommended by NICE as a second-line treatment for people with amRCC (TA178).[29]

Comparators

The comparators of interest and considered in this review were the interventions listed in the section *Interventions* and compared with each other or BSC. BSC in this context is defined as the standard care for people with RCC if the available drug therapies are contraindicated or not tolerated (e.g. social and palliative services, treatment for symptomatic relief). For the purposes of this review, we assumed that people randomised to a placebo group received BSC. In addition, studies were sought through the search

process covering a broader range of comparators to provide data to create additional links between the interventions in the mixed-treatment comparisons (MTCs). Full details of the additional comparators included in the searches along with the results are provided in *Chapter 3*.

Outcomes

The outcomes considered in this review are:

- overall survival (OS)
- progression-free survival (PFS)
- response rates
- AEs of treatment
- health-related quality of life (HRQoL).

The key outcomes for the primary analyses were OS and PFS and these were conducted using MTCs. Sensitivity analyses (SAs) for OS and PFS were conducted, which included data from observational studies, with a further analysis for PFS including studies that were rated as being at a high risk of bias. In addition, subgroup analyses were conducted based on MSKCC baseline prognostic score and number of prior TKIs, although data for these analyses were limited to very few interventions. MTCs were also used to analyse response rate data. Data for AEs and HRQoL were insufficient to allow meta-analysis and so they have been tabulated and discussed narratively in *Chapter 3, Assessment of effectiveness* and *Adverse events*.

Study design

The protocol and review set out to evaluate data from RCTs when available. The RCT data were not available to create a linked network between all of the interventions of interest in the review for the analyses of PFS and OS. Observational data were therefore sought in an attempt to identify outcome data to link the missing treatments (axitinib and sunitinib) into the MTCs. The nature of the observational studies is discussed in detail in *Chapter 3, Study characteristics*.

Overall aims and objectives of assessment

The objectives of this systematic review are to:

- evaluate the clinical effectiveness of axitinib, BSC, cabozantinib, everolimus, nivolumab and sunitinib in line with their respective marketing authorisations for amRCC that has been previously treated with a VEGF-targeted therapy
- evaluate the cost-effectiveness of axitinib, BSC, cabozantinib, everolimus, nivolumab and sunitinib in line with their respective marketing authorisations for amRCC that has been previously treated with a VEGF-targeted therapy
- identify key areas for further primary and secondary research.

The review focuses on patients who have received prior VEGF-targeted therapy because this is what clinical experts report as the expected first-line treatment for people in the UK with advanced RCC. This review does not cover the population of patients who have received only prior cytokines as these therapies are deemed to no longer be used routinely in UK clinical practice for the RCC population.

Chapter 3 Assessment of clinical effectiveness

Methods for reviewing effectiveness

Evidence on the clinical effectiveness of axitinib, cabozantinib, nivolumab, everolimus and sunitinib for people who have received previous VEGF-targeted therapy for the treatment of amRCC was identified by conducting a systematic review of the published research literature. The review was undertaken following the general principles published by the Centre for Reviews and Dissemination (CRD) and Cochrane.[40,41] The protocol for the systematic review is registered on PROSPERO (registration number CRD42016042384).[42]

This report contains reference to confidential information provided as part of the NICE appraisal process. This information has been removed from the report and the results, discussions and conclusions of the report do not include the confidential information. These sections are clearly marked in the report.

Identification of studies

To identify relevant studies, multiple electronic databases were searched:

- MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R)
- EMBASE
- The Cochrane Library [specifically Cochrane Database of Systematic Reviews, Cochrane Central Register of Controlled Trials (CENTRAL), Database of Abstracts of Reviews and Effects (DARE) and Health Technology Assessment (HTA) database].

Search strategies were designed to include medical subject headings (MeSH) and text terms for RCC, and the interventions of interest (axitinib, everolimus, nivolumab and sunitinib) with a RCT filter applied in MEDLINE and EMBASE searches. Additional search terms for interventions outside the scope of this report that could have been relevant for creating a connected network [e.g. temsirolimus (Torisel®, Pfizer, Kent, UK) and sorafenib] were also included in the original searches in January 2016 (see *Appendix 1*). The January 2016 searches also included search terms for cytokines. However, trials of interventions not listed in the final inclusion criteria were included in the final review only if they were needed to create a network linking the interventions and comparators listed in the final protocol for people with prior VEGF-targeted therapy.

It should be noted that cabozantinib was not included in the final protocol but was added to the review question in August 2016 in view of its introduction and potential availability in the UK. However, cabozantinib was not included in the electronic database searches because it was added to the review after these searches were run. Studies for cabozantinib were identified and validated via clinical experts, and the company submission for the NICE STA (TA463).[28]

Initial review of the identified RCTs from the electronic database searches revealed that there was a lack of suitable RCT data for axitinib and sunitinib to link them into a network for a MTC with the other interventions of interest. There was one RCT for axitinib compared with sorafenib (AXIS),[43] and no relevant RCTs including sunitinib. Sorafenib was not an intervention of interest and no other RCTs suitable for linking AXIS into a network were identified. As such, a decision was made in June 2016 to conduct further electronic database searches of MEDLINE and EMBASE specifically for observational studies of axitinib, sorafenib and sunitinib. The aim of these searches was to identify data for OS and PFS for axitinib and sunitinib to enable the inclusion of them in MTCs with the other interventions under review. Prospective and retrospective observational studies (matched control studies, case series, cohort and case–control studies) with a comparator group were sought and assessed for eligibility. A pragmatic decision was taken not to update the RCT searches in June 2016 as a result of time and resource constraints.

Search filters designed to retrieve reports by study design were identified via the InterTASC Information Specialists' Sub-Group search filter resource.[44] Filters developed and validated by the Scottish Intercollegiate Guidelines Network were used to identify RCTs and observational studies in MEDLINE and EMBASE. Search strategies were designed by a reviewer experienced in information retrieval and validated by a second reviewer experienced in designing search strategies, and terms for RCC and the interventions were tailored to the database searched.

Bibliographies of retrieved studies (RCTs, observational studies and systematic reviews) identified as relevant were manually reviewed for additional studies. Clinical trial registries (EU Clinical Trials Register and ClinicalTrials.gov) were also searched to identify planned or on-going clinical trials of interest. In addition, clinical experts were contacted with a request for information on any additional studies of which they had knowledge. Conference proceedings for the following conferences were also searched for further studies of potential relevance:

- European Multidisciplinary Meeting on Urological Cancers (EMUC), 2015
- EMUC, 2016
- European Cancer Congress, 2015
- European Society For Medical Oncology (ESMO), 2016
- American Society of Clinical Oncology (ASCO) Annual Meeting, 2015
- ASCO Annual Meeting, 2016
- American Society of Clinical Oncology-Genitourinary Cancers Symposium (ASCO-GU), 2015
- ASCO-GU, 2016.

No language or date restriction was applied to the searches. The electronic databases were searched from inception, with the initial search for RCT data carried out on 13 January 2016. Search results were uploaded into EndNote version X7.2 [Clarivate Analytics (formerly Thomson Reuters), Philadelphia, PA, USA] and deduplicated. Electronic database searches for observational studies were carried out in June 2016. Full details of the search strategies are presented in *Appendix 1*.

Two researchers [Charlotta Karner (CK) and one of Natalie Masento (NM), George Osei-Assibey (GOA) or Claire Fiatikoski (CF)] independently screened titles and abstracts, initially for RCTs and systematic reviews for eligibility. Full texts were retrieved and appraised (CK and either GOA or NM) for publications agreed to be potentially relevant and those for which consensus could not be reached on the basis of the abstract alone. Discrepancies were resolved by discussion, with involvement of a third reviewer [Victoria Wakefield (VW)] if consensus could not be reached. After appraisal of full text publications for RCTs, study type eligibility was broadened to connect all treatments in the network, and abstracts were reappraised by two reviewers (CK and NM) for comparative observational studies of axitinib, sorafenib or sunitinib. The search results from the June 2016 searches for observational studies were also appraised by CK and NM following deduplication by VW. Full-text papers of potentially eligible studies were then ordered and appraised independently (CK and NM).

Inclusion and exclusion criteria

Eligibility criteria for the review of clinical effectiveness in people who have received previous VEGF-targeted therapy were as specified in the final protocol, with the exception of cabozantinib, which was added in August 2016 (summarised in *Table 2*). The interventions of interest were axitinib, cabozantinib, everolimus, nivolumab and sunitinib. The review included RCTs of any intervention of interest along with comparative observational studies of axitinib, sorafenib and sunitinib compared with any intervention of interest. Pre-clinical studies and those conducted in animals, narrative reviews, editorials, opinions, case reports and systematic reviews were excluded from the review. Studies were included if the treatments were evaluated in a population with prior VEGF-targeted therapy for RCC and were compared with each other, placebo or BSC. Studies were excluded if none of the outcomes of interest was reported. Observational studies were included only if they reported PFS or OS data in a way that could be incorporated into the MTC [i.e. as a

hazard ratio (HR) or Kaplan–Meier (KM) curve with the number of patients at risk]. Studies of sorafenib compared with any of the interventions of interest were included to enable the inclusion of the AXIS trial for axitinib in the MTCs.[43]

Data abstraction

Data were extracted independently by two reviewers (VW and GOA) into a standardised data extraction form in Microsoft Word (Microsoft Corporation, Redmond, WA, USA) for three studies to pilot the suitability of the data extraction form. Subsequently, two reviewers (CF, CK, GOA, NM or VW) independently extracted data for each of the remaining studies into a modified data extraction form, with validation of the data by a third reviewer (CK or VW). Information extracted included details on study design and methodology, the baseline characteristics of the study population and data on outcomes of interest. A pragmatic decision was made to restrict the extraction of AEs of treatment (AE) data to those relating to common terminology criteria for adverse events (CTCAE), v3.0 or later, ≥ grade 3 AEs owing to the large number of AE data potentially reported.[45] Discrepancies in data extraction forms were resolved by discussion, with the involvement of a fourth reviewer (CK or VW, depending on who was the third reviewer) when necessary. Data extraction forms for the included studies are provided in *Appendix 8*.

Critical appraisal strategy

Two reviewers [CK, Kayleigh Kew (KK), NM or VW] independently assessed the quality of the clinical effectiveness studies. Discrepancies were resolved by discussion, with involvement of a third reviewer (CK or VW, dependent on who was the third reviewer) when necessary. Study quality was assessed according to recommendations by the CRD and the *Cochrane Handbook for Systematic Reviews of Interventions*.[40,41] Study quality for RCTs was recorded using the Cochrane Risk of Bias Tool[46] and was reported in tables for each study (see *Appendix 9*). Study quality for the non-randomised studies was assessed using the Risk Of Bias In Non-randomised Studies – of Interventions (ROBINS-I) tool[47] (see *Appendix 9*).

Outcome-specific risk of bias was determined for the outcomes for which data were extracted. A total of three bias assessment categories were used for RCTs: low, unclear and high. Within a study, outcome data were rated as being at a low risk of bias when all domains were associated with a low risk of bias, at an unclear risk of bias when one or more domains had an unclear risk of bias, and at a high risk of bias when one or more domains was rated as being at a high risk of bias. Observational studies were assessed for bias using the following five categories: no information, low, moderate, serious and critical. Similar to the overall bias assessment rating for RCTs, observational studies were rated as being at the highest bias rating that they received for any individual domain. The bias severity for observational studies ascended from low to critical and was also assessed and graded for each outcome.

Methods of data synthesis

Details of the results on clinical effectiveness and quality assessment for each included study are presented in structured tables and an overall assessment of study quality is provided as a narrative summary (see *Quality assessment of studies*). The possible effects of study quality on the interpretation of clinical effectiveness data and review findings are discussed, when relevant.

The analysis of clinical effectiveness was based on intention-to-treat (ITT) populations when possible. ITT was defined as people being analysed in the treatment group to which they were allocated at randomisation irrespective of whether they changed treatment, withdrew or were lost to follow-up. Pairwise meta-analysis was not possible owing to the absence of more than one RCT per pairwise comparison of interest. The comparative clinical effectiveness of interventions was investigated instead via a MTC. The methods used for the MTC followed the guidance described in the NICE Decisions Support Unit's Technical Support Documents for Evidence Synthesis.[48,49] MTCs were conducted using a Bayesian Markov chain Monte Carlo simulation in WinBUGS version 1.4 (MRC Biostatistics Unit, Cambridge, UK). This has the additional advantage of being

able to calculate direct probability statements for which treatment is the most effective, even when standard methods might determine no significant difference between treatments.[48,50–52] The following were implemented for each analysis.

- Uniform priors (also called 'uninformed' or 'flat' priors) were used.
- All outcomes were considered independent.
- To ensure convergence on the posterior distribution results for all clinical effectiveness outcomes analysed were based on a 'burn in' of a minimum of 10,000 iterations.
- HR was used as the summary effect estimate for PFS and OS.
- Odds ratio (OR) was used as the summary effect estimate for response rate.
- Alongside HRs and ORs, 95% credible intervals (CrIs) were reported (a 95% CrI that does not cross a value of 1 is analogous to a statistically significant difference at the 5% level of significance).

Fixed-effects and random-effects models were explored. However, as typically only one trial informed each pairwise comparison, a pragmatic decision was made to use the fixed-effects model for all outcomes. This decision was supported by the impact of using an uninformed prior for the between trial heterogeneity in a random-effects model. The prior 'overwhelmed' the influence of the available data for analysis with the posterior estimation of tau approximating the prior value used.

Sensitivity analyses were carried out for the outcomes of OS and PFS. SAs included observational studies, and a second SA that included studies that were rated as being at an overall critical risk of bias was conducted for PFS. No studies were rated as being at a critical risk of bias for OS.

Inconsistency in the MTC networks was assessed when loops were present allowing a comparison of the direct and indirect effect estimates. However, this was only possible in the SAs for PFS and OS.

Subgroup analyses were carried out as planned for both PFS and OS, based on:

- the number of prior therapies
- baseline prognostic score (e.g. MSKCC).

No assessment of publication bias was conducted as a result of the limited number of studies identified for each intervention.

Results

Quantity of research available

As discussed in *Identification of studies*, the electronic database searches were conducted in two parts: the first for RCTs and the second for observational studies. Results from both searches were screened for both RCT and observational studies meeting the review inclusion criteria detailed in *Inclusion and exclusion criteria*. The PRISMA diagram in *Figure 1* provides a summary of the search results for both sets of searches.

A total of 6079 records were identified after deduplication of the electronic database search results for the searches conducted in January 2016. Full-text papers for 112 articles were assessed and, of these, 88 articles were excluded for reasons including not an intervention of interest, no suitable outcome data and incorrect study population. The 24 included publications related to three RCTs (20 publications) and two crossover studies (four publications).[43,53–56] A further RCT (three publications) for cabozantinib,[57] an intervention added to the review following the search date, was identified via clinical experts and the related company submission for the NICE STA (TA463).[28]

The electronic database searches conducted in June 2016 for observational studies of axitinib, sorafenib and sunitinib identified 1120 records following deduplication. Title and abstract appraisal led to the

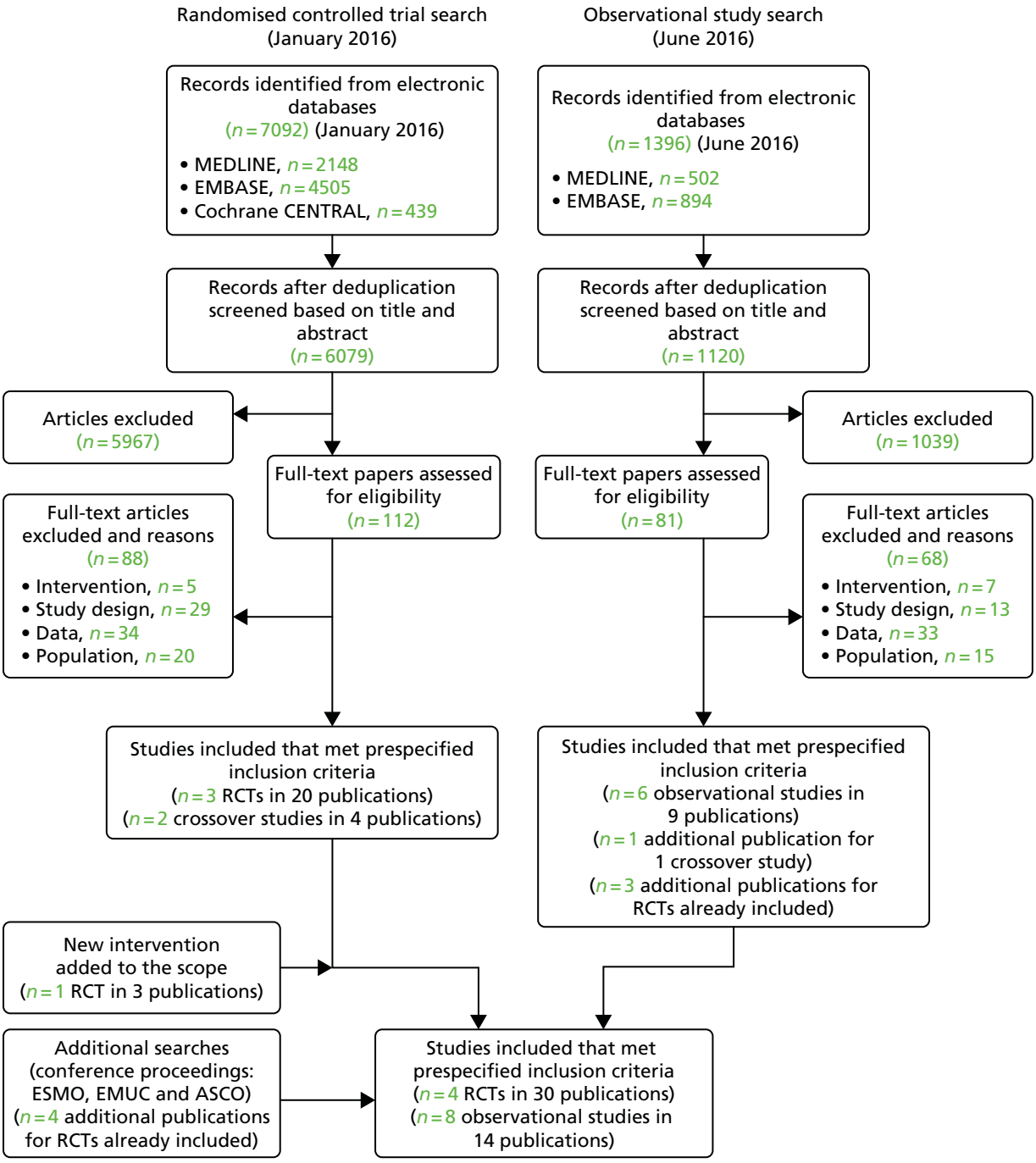


FIGURE 1 The PRISMA flow diagram for search results.

exclusion of 1039 articles, leaving 81 publications for full-text appraisal. The 81 full-text papers screened for potential inclusion resulted in 68 of these subsequently being excluded. The most common reason for exclusion at full-text appraisal was a result of not reporting outcome data of interest ($n = 33$). The 13 final included publications from this search related to six observational studies (nine publications) along with one additional publication for an already included crossover study[58–63] and three additional publications for RCTs already included.

Searches of the conference abstracts from ESMO, EMUC, ASCO and ASCO-GU resulted in the inclusion of a further four publications for two already included RCTs.[54,57] Searches of ClinicalTrials.gov and the EU Clinical Trials Register conducted in February 2017 identified a total of 338 records, of which 72 were

duplicates. Nine records were identified as potentially relevant but, on closer inspection, did not meet the eligibility criteria for the review.

In summary, a total of four RCTs (in 30 publications) and eight observational studies including two crossover studies (14 publications) met the inclusion criteria.[43,53–63] A list of the included studies and their associated publications can be found in *Appendix 3*. The included studies and their findings are discussed further in sections *Study characteristics* to *Summary of the results of the review of clinical effectiveness*. A list of publications screened but subsequently excluded (with reasons for exclusion) from the review is available in *Appendix 10*.

Study characteristics

The key characteristics of the 12 included studies are summarised in *Table 3*. As discussed in *Quantity of research available*, four of these studies were RCTs (AXIS,[43] CheckMate 025,[54] METEOR[57] and RECORD-1[53]) and the remaining eight were observational studies.[55,56,58–63] All four of the RCTs[43,53,54,57] were multicentre, international, Phase III clinical trials. RECORD-1[53] was the only double-blind study. The remaining three RCTs (AXIS[43], CheckMate 025[54] and METEOR[57]) were open-label designs and some of the outcomes, such as PFS, were assessed via blinded independent review panels. This is discussed in more detail along with the quality assessment in *Quality assessment of studies*.

The eight observational studies comprised the post-crossover part of two RCTs[55,56] and six retrospective cohort studies.[58–63] The two crossover RCTs[55,56] recruited patients who were treatment naive, defined as having received no prior systemic therapy and, thus, they did not meet our eligibility criteria. The patients in both ESPN[55] and SWITCH[56] were randomised to receive a VEGF-targeted therapy and following disease progression or discontinuation from randomised study treatment; they were then eligible to crossover and receive the alternative study drug. As such, only the data from this 'post-crossover' period meet the inclusion criteria for this review, because all of the patients were then pre-treated. The patients were not randomised to the second treatment in these crossover studies and so the data from the post-crossover period has been treated as observational data. The six remaining observational studies[58–63] were retrospective studies that include medical chart and note reviews. Similar to the crossover RCTs, only data for the second period of a treatment sequence could be included from three of the retrospective studies.[58,60,61] These three studies reviewed data from patients who had received either sunitinib followed by sorafenib or vice versa, and so only data for the second period were from a population pre-treated with a VEGF-targeted therapy. Further details relating to the study design and risk of bias are discussed in *Quality assessment of studies*.

Population

The population in all of the studies comprised patients with amRCC who had received at least one prior VEGF-targeted therapy apart from AXIS, in which only a subgroup of 54% of patients received prior VEGF-targeted therapy.[43] Data for the eligible subgroup were used for the primary outcome analyses (OS and PFS) to minimise potential bias, but data used for the secondary outcomes refer to the full AXIS population.[43] The sample size varied across the studies from 33[58] to 821.[54] The sample size was generally higher among the RCTs (range 416[64] to 821[54]) than among the observational studies (range 33[58] to 452[62]).

All of the studies recruited adults (people aged ≥ 18 years), with the median age of patients across the studies, when age was reported at baseline, generally between 60 and 70 years. When ethnicity was reported, > 70% of the patients in the studies were white and there was a higher proportion of males than females in all studies (lowest proportion 58% in the everolimus group in the ESPN study[55] and highest proportion 86% in both groups in Paglino *et al*.[60]).

Seven out of the eight studies that did describe RCC subtype recruited a solely or primarily clear-cell population. ESPN was the only study to recruit a non-clear-cell population, or clear cell with at least 20% sarcomatoid features.[55] The type of RCC was not reported in the three retrospective studies comparing the sequence of sunitinib and sorafenib,[58,60,61] or in SWITCH.[56] The stage of RCC was generally poorly reported,

TABLE 3 Summary table of included studies

	Key characteristic													
									ECOG (%)		MSKCC (%)			Treatment duration, months (follow-up)
Study	Study design	Treatments	Sample size	Type	Prior therapies	Age (years)	Male (%)	Ethnicity white (%)	0/1	2	F	I	P	
AXIS[43]	Phase III open-label RCT	Axitinib	361	CC	1 prior TKI; other prior therapies permitted	61	73	77	99	1	28	37	33	8.2 (NR)
		Sorafenib	362			61	71	74	100	0	28	36	33	5.2 (NR)
Calvani *et al.*, 2013[58]	Retrospective observational	Sunitinib	15	NR	1 prior TKI; other prior therapies permitted	70[a]	80	NR	93	7	20	73	7	NR (NR)
		Sorafenib	18			61[a]	61		78	22	22	78	0	
CheckMate 025[54]	Phase III open-label RCT	Nivolumab	410	CC	1/2 prior antiangiogenic; no prior mTORi	62	77	86	NR		35	49	16	5.5 (NR)
		Everolimus	411			62	74	89			36	49	15	3.7 (NR)
ESPN[55]	Phase II crossover RCT	Sunitinib	21	NCC	1 prior mTORi	58	69	80	100	0	11	83	6	NR (23.6)
		Everolimus	23		1 prior TKI	60	58	76	100	0	12	88	0	
Iacovelli *et al.*, 2015[59]	Retrospective observational	Sorafenib	90	CC	2 prior targeted therapies (TKI or other)	63	74	NR	81	19	NR			NR (NR)
		Everolimus	143											
METEOR[57]	Phase III open-label RCT	Cabozantinib	330	CC	1 or more prior TKIs; no prior mTORi	63[a]	77	82	100	0	45	42	12	8.3 (18.7)
		Everolimus	328			62[a]	73	80	100	0	46	41	13	4.4 (18.8)
Paglino *et al.*, 2013[60]	Retrospective observational	Sunitinib	26	NR	1 prior TKI (sorafenib or sunitinib) and mTORi	61	86	NR	96	4	46	54	0	NR (NR)
		Sorafenib	14			63	86		93	7	57	29	14	
Porta *et al.*, 2011[61]	Retrospective observational	Sunitinib	90	NR	1 prior TKI (sorafenib)	58	82	NR	98	2	50	39	10	NR (NR)
		Sorafenib	99		1 prior TKI (sunitinib)	60	68		97	3	41	26	32	

continued

TABLE 3 Summary table of included studies (*continued*)

	Key characteristic													
									ECOG (%)		MSKCC (%)			Treatment duration, months (follow-up)
Study	Study design	Treatments	Sample size	Type	Prior therapies	Age (years)	Male (%)	Ethnicity white (%)	0/1	2	F	I	P	
RECORD-1[53]	Phase III double blind RCT	Everolimus	277	CC	1/2 prior TKI; bevacizumab and cytokines permitted, no mTORi	61[a]	78	NR	NR		29	56	15	4.6 (NR)
		BSC	139			60[a]	76				28	57	15	1.9 (NR)
SWITCH[56]	Phase III crossover RCT	Sunitinib	103	NR	1 prior TKI; no other prior systemic therapy	62	79	NR	99	0	46	54	0	6.4 (10.3[a])
		Sorafenib	76			63	74		100	0	50	50	0	5.9 (10.3)
Vogelzang *et al.*, 2016[62]	Retrospective observational	Everolimus	325	85% CC	1 prior TKI; no prior cytokines	61[a]	70	NR	80	19	NR			NR (15[a])
		Axitinib	127			60[a]	65		84	16				NR (13[a])
Wong *et al.*, 2014[63]	Retrospective observational	Everolimus	233	91% CC	1 prior TKI; no mTORi, cytokine, bevacizumab	64	70	82	NR		NR			NR (12.9)
		Sorafenib	123			66	72	79						NR (12.1)

CC, clear-cell variant; F, favourable MSKCC category; I, intermediate MSKCC category; P, poor MSKCC category; mRCC, metastatic renal cell carcinoma; NCC, non-clear-cell variant; NR, not reported.

a Mean values for which median was not reported.

Notes

For ECOG and MSKCC, percentages that do not total 100 are due to missing data.

although all studies had inclusion criteria that patients were required to have amRCC. This inclusion criteria would suggest that all patients were a minimum of stage II, although the spread of patients over the different stages across the studies is unclear. Baseline prognostic score was not reported for all studies although, when reported, nearly all patients had a reasonably good performance status (i.e. ECOG performance status 0 or 1).

Study inclusion and exclusion criteria with regard to prior therapies varied across the studies (see *Table 3*). Prior therapies in the table reflect the population included in this review and not necessarily the study inclusion criteria (e.g. treatment-naive patients were recruited at the start of ESPN,[55] but had all received one prior treatment in the period included in our analyses). Eight study populations had received one prior TKI treatment (or mTORi in the case of ESPN),[43,55,56,58,60–63] and four included patients who had received two lines of prior therapy: RECORD-1[64] specified one or two prior TKIs, CheckMate 025[54] and Iacovelli *et al.*[59] allowed two prior targeted therapies, and METEOR[57] allowed any number of prior TKI treatments. Other types of prior therapy (e.g. chemotherapy, cytokines, bevacizumab) were allowed in most studies, and prior mTORi therapy was usually not permitted. Prior cytokine use was exclusionary in Vogelzang *et al.*[62] and Wong *et al.*[63] The impact of prior therapies is discussed and explored through subgroup analyses (see *Subgroup analyses*).

Formal statistical tests for between-treatment group differences at baseline were generally not reported. On visual inspection, the RCT populations appear to be balanced but there were notable differences between treatment groups in some of the observational studies. Imbalances include the percentage of second-line patients in Porta *et al.*[61] (62% sunitinib vs. 29% sorafenib); mean age in Calvani *et al.*[58] (70 years sunitinib vs. 61 years sorafenib); percentage male imbalances in Calvani *et al.*,[58] ESPN,[55] and Porta *et al.*;[61] and differences in the distribution of MSKCC prognostic scores in Paglino *et al.*[60] and Porta *et al.*[61] Therefore, the results from the observational studies may be subject to bias when the analyses have not been adjusted for these baseline imbalances (e.g. age could have a substantial impact on OS irrespective of treatment effect).

Intervention and comparator

The RCTs evaluating all of the treatments of interest apart from sunitinib were identified. The inclusion of observational studies on axitinib, sorafenib and sunitinib also led to the inclusion of observational studies for everolimus. There were two studies that included axitinib (one RCT[43] and one observational study),[62] one study for cabozantinib (one RCT),[57] seven studies for everolimus (three RCTs and four observational studies),[53–55,59,62,63] one study for nivolumab (one RCT[54]) and five studies for sunitinib (five observational studies).[55,56,58,60,61] There were seven studies that also included sorafenib (one RCT,[43] and six observational studies),[56,58–61,63] and one that included BSC (one RCT).[53] Network diagrams for PFS (*Figures 2* and *3*) and OS (*Figures 4* and *5*) illustrate which direct treatment comparisons contributed to each MTC.

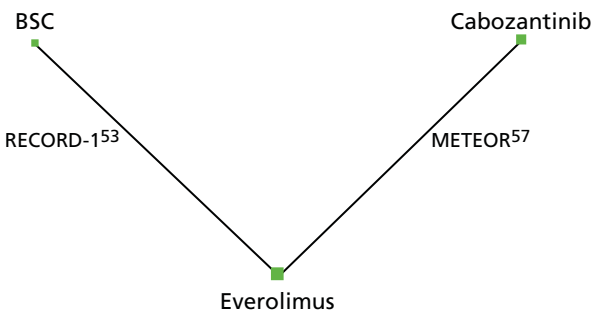


FIGURE 2 Network diagram for PFS (primary analysis). Notes: the size of the nodes represent the number of patients on each intervention. The thickness of the lines represents the number of studies informing the direct comparison.

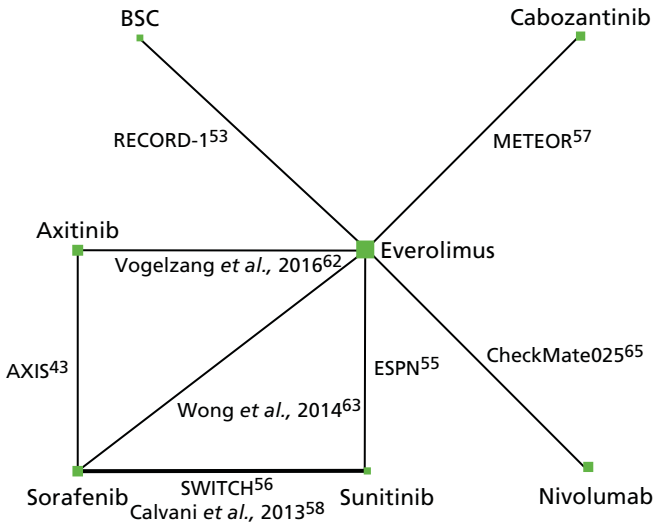


FIGURE 3 Network diagram for PFS for SA1. Notes: the size of the nodes represent the number of patients on each intervention. The thickness of the lines represents the number of studies informing the direct comparison.

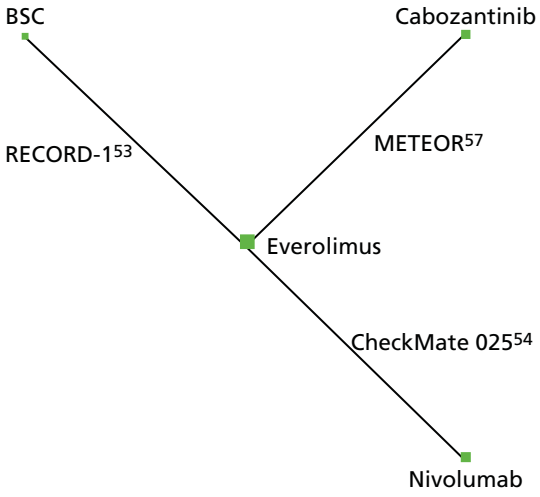


FIGURE 4 Network diagram for OS primary analysis. Notes: the size of the nodes represent the number of patients on each intervention. The thickness of the lines represents the number of studies informing the direct comparison.

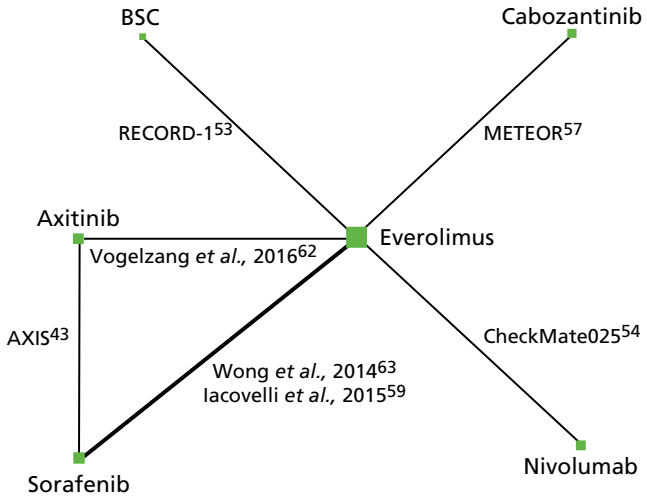


FIGURE 5 Network diagram for OS SA. Notes: the size of the nodes represent the number of patients on each intervention. The thickness of the lines represents the number of studies informing the direct comparison.

The active study drug dose was not specified in all studies but, when it was specified, it was the standard licensed dose (see *Chapter 1, Description of the technologies under assessment*). Doses were varied according to clinician and patient factors, with limited details reported on these dose adjustments. No study reported the explicit use of any concomitant medications.

The duration of treatment was only reported in the four RCTs and one of the crossover observational studies.[43,53,54,56,57] As such, we are uncertain of the variation in treatment duration within the observational studies and whether or not it differed between them and the RCTs. The median treatment duration in the five studies, when it was reported, varied from 1.9 months [placebo (BSC) group of RECORD-1] to 8.3 months (cabozantinib group of METEOR).[57,64] Treatment was reported in the RCTs to be continued until disease progression, unacceptable toxicity or withdrawal of consent. Treatment discontinuations are discussed alongside the quality assessment in *Quality assessment of studies*. Median length of follow-up was also poorly reported among the included studies with only four studies reporting data.[55,57,62,63] The median length or follow-up ranged from 12.1 months to 23.6 months.[55,63] One study reported mean length of follow-up, which was 10.3 months.[56] Study duration was also not reported for all studies and varied widely when it was reported (range 22–44 months), although the longer studies relate to the crossover RCTs used as observational studies (and so the first part of the study is not relevant to this review). These data for length of study and length of follow-up should, thus, be interpreted with caution.

Details on subsequent therapies following treatment discontinuation in the included studies was limited. Most studies allowed subsequent treatment in the event of progression or intolerable toxicity on the study drug, but gave minimal detail about what was actually received and the possible impact it might have on the results. However, in RECORD-1, BSC (placebo) patients could cross over to receive open-label everolimus during the study and it was reported that 76.2% of patients did so.[53] The OS results of RECORD-1 used in the MTC are crossover adjusted in an attempt to address the potential bias introduced from the high level of crossover.[53,65] Treatment crossover was not reported to have occurred in any other studies, except when it was part of the study design. In METEOR and Checkmate 025, patients were allowed to continue the study therapy after initial disease progression if a clinical benefit was noted by the investigator.[54,57]

Outcomes

The outcomes of interest to this review and reported in the included studies are listed in *Table 4*. Data for all of the outcomes were available, although outcome data were not available for all of the interventions. The primary outcome in the majority of the studies was PFS. The data reported in the six retrospective observational studies comprised only PFS and OS data, which as discussed earlier (see *Methods for reviewing effectiveness*) have only been used in SAs.[58–63] Data from the RCTs also included response rate, HRQoL (RCT data only) and AE data; however, the reporting of these outcomes varied across the studies both in terms of their presence and the type of data (e.g. the HRQoL tools used and individual AEs reported varied across studies).[43,54–57,64] The available data will be discussed in detail in the results subsections below (see *Assessment of effectiveness* to *Subgroup analyses*).

Quality assessment of studies

The four RCTs were of good methodological quality (*Table 5*).[43,53,54,57] Using the Cochrane Risk of Bias tool,[46] all were rated as being at a low risk of selection biases (random sequence generation and allocation concealment) because they used computerised code generators and automated allocation systems to assign participants to groups. AXIS,[43] CheckMate 025[54] and METEOR[57] were open-label head-to-head comparisons and are, thus, considered to be at a high risk of bias for blinding of participants and personnel. RECORD-1 was a placebo-controlled double-blind trial and so was rated as being at a low risk of bias for this domain.[53] Risks of bias for detection (blinding of outcome assessors), attrition (incomplete outcome data), reporting (selective outcome reporting) and other biases varied between studies and within studies by outcome. Hence, these bias domains were assessed for each outcome separately and have been described in more detail in the assessment of effectiveness section (see *Assessment of effectiveness*). When possible, risk-of-bias judgements refer to the particular outcome data included in the analysis and this is noted in each study's quality assessment (see *Appendix 9*). In general, OS and PFS are considered to be at a low risk of detection

TABLE 4 Summary of outcome data included from the included studies

	Outcome								
	Primary outcome		Primary outcome subgroup analyses				Secondary outcome		
Study	PFS	OS	PFS by prior treatment	PFS by prognostic score	OS by prior treatment	OS by prognostic score	ORR	HRQoL (narrative)	AE (narrative)
AXIS[43]	✓[a]	✓[a]	✓		✓			✓	✓
Calvani *et al.*, 2013[58]	✓[a]								
Checkmate-025[54]		✓			✓		✓	✓	✓
ESPN[55]	✓[a]								
Iacovelli *et al.*, 2015[59]		✓[a]							
METEOR[57]	✓	✓	✓	✓	✓	✓	✓	✓	✓
Paglino *et al.*, 2013[60]	✓[a]								
Porta *et al.*, 2011[61]	✓[a]		✓	✓					
RECORD-1[53]	✓	✓	✓	✓			✓	✓	✓
SWITCH[56]	✓[a]								
Vogelzang *et al.*, 2014[62]	✓[a]	✓[a]		✓		✓			
Wong *et al.*, 2014[63]	✓[a]	✓[a]							

ORR, objective response rate.
a Data included in a SA, not in primary MTC.

TABLE 5 Summary of Cochrane risk-of-bias assessment for RCTs

	Study			
Criteria	AXIS[43]	Checkmate-025[54]	METEOR[57]	RECORD-1[53]
General risk of bias				
Sources of bias related to study characteristics				
Random sequence allocation	✓	✓	✓	✓
Allocation concealment	✓	✓	✓	✓
Blinding: participant and personnel	✗	✗	✗	✓
Outcome specific				
PFS				
Blinding: outcome assessment	✓	✗	✓	✓
Incomplete outcome data	✓	✓	?	✓
Selective reporting	✓	✓	✓	✓
Other biases	?	?	N/A	?
Overall survival				
Blinding: outcome assessment	✓	✓	✓	✓
Incomplete outcome data	✓	?	?	✓
Selective reporting	✓	✓	✓	✓
Other biases	✓	✓	?	?

TABLE 5 Summary of Cochrane risk-of-bias assessment for RCTs (*continued*)

Criteria	Study			
	AXIS[43]	Checkmate-025[54]	METEOR[57]	RECORD-1[53]
Response rate				
Blinding: outcome assessment	✓	✗	✓	✓
Incomplete outcome data	?	✓	?	✓
Selective reporting	✓	✓	✓	?
Other biases	N/A	N/A	N/A	?
AEs				
Blinding: outcome assessment	✗	✗	✗	✓
Incomplete outcome data	✓	✓	✓	✓
Selective reporting	✓	✓	✓	✓
Other biases	N/A	?	N/A	N/A
HRQoL				
Blinding: outcome assessment	✗	✗	✗	✗
Incomplete outcome data	✓	✓	?	✓
Selective reporting	✓	✓	✓	✗
Other biases	N/A	N/A	N/A	N/A

✗, high risk; ✓, low risk; N/A, not applicable; ?, unclear risk.

and reporting biases for all RCTs, except for a high risk of PFS detection bias in CheckMate 025 because the end point was not assigned by an independent review committee.[54] None of the outcomes in the RCTs was rated as being at a high risk of attrition bias; all used appropriate censoring for the time-to-event analyses, although there is some uncertainty regarding immature data, particularly for OS in CheckMate 025 and METEOR.[54,57] The RCTs were mostly rated as being at a low or unclear risk of reporting biases. Other possible sources of bias recorded mainly pertain to group differences in the rate and type of subsequent therapies received, and in the way drug dose could be managed.

The non-RCTs included in the OS and PFS SAs, including retrospective cohorts and crossover RCTs, for which only the crossed-over phase met the inclusion criteria for the review, were of low methodological quality. Overall, ROBINS-I ratings were serious or critical across studies and outcomes (*Table 6*). The risk of bias due to confounding was mostly serious in four studies[55,56,58,59] because key variables identified in our protocol were not adjusted for in the analyses. Paglino *et al.*[60] and Porta *et al.*,[61] both reporting PFS, were rated critical for confounding bias because analyses were not adjusted for significant baseline imbalances; these two studies have thus been omitted from the PFS SA to explore results including non-RCT data. There were no other critical ratings across other domains. Vogelzang *et al.*[62] and Wong *et al.*,[63] both reporting PFS and OS, are rated as being at a moderate risk of confounding bias because results were adjusted for some but not all key confounding domains. The non-RCTs are generally rated as being at a serious risk of selection biases as a result of their primarily retrospective designs; ESPN and SWITCH are at a moderate risk because participants were randomised to the first phase of the study.[55,56] There were no concerns in any of the studies regarding classification of the interventions and deviation from the intended interventions. Bias due to missing data varied across studies, ranging from low[56,59,62] to serious (Wong *et al.*,[63] PFS and OS); all studies censored patients who did not progress or, at last contact, were lost to follow-up but Wong *et al.*[63] excluded participants if there were any missing baseline data needed for the multivariate analyses. Bias in measurement of the outcome was low risk for OS and serious for all studies reporting PFS with the exception of ESPN, which was the only non-RCT to use an independent review panel.[55] The use of Response Evaluation

TABLE 6 Summary of ROBINS-I risk-of-bias assessments in non-randomised studies

	Study									
	Calvani *et al.*, 2013[58]	ESPN[55]	Iacovelli *et al.*, 2015[59]	Paglino *et al.*, 2013[60]	Porta *et al.*, 2011[61]	SWITCH[56]	Vogelzang *et al.*, 2014[62]		Wong *et al.*, 2014[63]	
Outcome	PFS	PFS	OS	PFS	PFS	PFS	PFS	OS	PFS	OS
Confounding	✗	✗	✗	✗✗	✗✗	✗	~	~	~	~
Selection	✗	~	✗	✗	✗	~	✗	✗	✗	✗
Intervention classification	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Intervention deviations	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Missing data	~	✓	✓	NI	NI	✓	✓	✓	✗	✗
Outcome measures	✗	✓	✓	✗	✗	✗	✗	✓	✗	✓
Outcome reporting	~	✓	✓	~	~	~	✗	✓	✗	✓
Overall judgement	✗	✗	✗	✗✗	✗✗	✗	✗	✗	✗	✗

✗✗, critical risk; ✓, low risk; ~, moderate risk; NI, no information; ✗, serious risk.

Criteria In Solid Tumours (RECIST) criteria to assign progression for PFS reduced the risk of bias in the selection of the reported result by preventing PFS being measured in multiple ways,[56,58,60,61] but studies generally did not prespecify how analyses would be undertaken. Studies not using RECIST criteria were rated as being at a serious risk of bias for this domain.[62,63]

Assessment of effectiveness

Progression-free survival

Comparative clinical effectiveness of PFS was evaluated through a MTC. The primary network generated comprised just two studies (RECORD-1 and METEOR) and provides information on three treatments: cabozantinib, everolimus and BSC (see *Figure 2*).[53,57] As described in *Chapter 2, Comparators*, the term BSC has been used to refer to placebo throughout this report and placebo is assumed to be a surrogate for BSC. In RECORD-1, PFS was defined as the time from randomisation to the first documentation of disease progression or death (from any cause) assessed via blinded independent central review. Similarly, in METEOR, PFS was defined as the time from randomisation to radiographic progression per RECIST or death from any cause. In CheckMate 025,[54] PFS was also assessed using the RECIST criteria, which it has been suggested does not take into account 'tumour flare'. Tumour flare is a result of the immune response to immunotherapies like nivolumab and may be misinterpreted as progression. However, while the evidence review groups' (ERGs) clinical experts consider that, in theory, the RECIST criteria may be conservative for assessing PFS in patients treated with immunotherapies like nivolumab, they also consider tumour flare to be rare in clinical practice. Unfortunately, no estimates of PFS for nivolumab could be generated using MTC because the KM curves suggested that proportional hazards (PHs) did not hold for this outcome in CheckMate 025,[54] which was the only study evaluating nivolumab in this review. Axitinib and sunitinib were not included in the primary analysis as these interventions were assessed in studies that could not be connected in a network of high-quality studies. SA1 included observational studies of reasonable quality in addition to the RCTs; SA2 included all relevant studies identified (i.e. RCTs and observational studies of any quality, including studies deemed to be at a critical risk of bias).

Sensitivity analysis 1 incorporated the two RCTs from the primary analysis and five observational studies.[55,56,58,62,63] The network created by these additional studies facilitated the inclusion of one additional RCT (AXIS – prior sunitinib subgroup) and provides PFS estimates for axitinib and sunitinib.[43] A network diagram for SA1 is provided in *Figure 3*.

Sensitivity analysis 2 included the eight studies in SA1 as well as the two observational studies[60,61] rated being at a critical risk of bias because of confounding from significant imbalances of patient characteristics at baseline.[60,61] The two additional studies in SA2[60,61] provided additional data on sunitinib and sorafenib with the network otherwise remaining the same as for SA1.

Everolimus was chosen as the baseline treatment for the MTCs because of the comparatively large number of studies available for analysis.

The results of the primary analysis for PFS are presented in *Table 7*. Cabozantinib (HR 0.17, 95% CrI 0.12 to 0.24) and everolimus (HR 0.33, 95% CrI 0.25 to 0.43) showed a statistically significant PFS benefit compared with BSC, and cabozantinib showed a benefit over everolimus (HR 0.51, 95% CrI 0.41 to 0.63).

The results of SA1 were consistent with that of the primary analysis and provided estimates for additional treatment comparisons (see *Appendix 4*). Everolimus (HR 0.33, 95% CrI 0.25 to 0.43), cabozantinib (HR 0.17, 95% CrI 0.12 to 0.24), axitinib (HR 0.31, 95% CrI 0.214 to 0.44) and sunitinib (HR 0.27, 95% CrI 0.17 to 0.40) all showed a benefit on PFS compared with BSC. Cabozantinib has significantly better PFS than all other treatments: everolimus (HR 0.51, 95% CrI 0.41 to 0.63), sunitinib (HR 0.63, 95% CrI 0.44 to 0.95), axitinib (HR 0.54, 95% CrI 0.40 to 0.76) and BSC (HR 0.17, 95% CrI 0.12 to 0.24). Differences in PFS between sunitinib, everolimus and axitinib were not statistically significant. Based on sampling from the MTC, cabozantinib has a 99% probability of being the most effective treatment for improving PFS compared with the other treatments included in the analysis.

The results of SA2 (see *Appendix 4*) were similar to those of SA1 and the primary analysis, with no change in the statistical significance of any of the results.

The residual deviance was similar to the number of unconstrained data points in all of the MTC analyses for PFS except for SA2, for which the residual deviance was slightly lower than the number of unconstrained data points (8 points vs. 10 points, respectively for SA2). These results suggest that the fixed-effects MTC model was a good fit for the primary analysis and SA1. The random-effects model was deemed unsuitable as discussed in *Methods of data synthesis*.

The direct and indirect estimates of the HRs generated for the interventions in the connected loops in SA1 were compared to assess possible inconsistency in the MTC. There were two loops in SA1: loop 1, consisting of everolimus, axitinib and sorafenib, and loop 2, consisting of everolimus, sorafenib and sunitinib. The results of the inconsistency assessments demonstrated no evidence of significant inconsistency ($p < 0.05$, see *Appendix 5*).

TABLE 7 The PFS primary analysis. HRs and associated CrIs; HR < 1 favours treatment in the left-hand column

	Treatment		
Treatment	BSC	Cabozantinib	Everolimus
Everolimus	0.33 (0.25 to 0.43)	1.95 (1.59 to 2.42)	–
Cabozantinib	0.17 (0.12 to 0.24)	–	0.51 (0.41 to 0.63)
BSC	–	6.04 (4.24 to 8.40)	3.06 (2.31 to 3.98)

Cells highlighted in green indicate statistically significant results.

Overall survival

The primary analysis for OS included three RCTs[53,54,57] covering four interventions: cabozantinib, everolimus, nivolumab and BSC. CheckMate 025 was included in the MTC analysis for OS because inspection of the KM curves suggested that the assumption of PHs holds from 6 weeks onwards.[54] No high-quality studies connect axitinib or sunitinib to the network for the primary analysis. As 6 weeks is a small proportion of time in the analysis of OS, the pragmatic decision was made to include CheckMate 025 in the MTC, particularly given the absence of alternative sources of data for nivolumab.[54] The data in the MTC analysis for OS from RECORD-1 was adjusted for crossover using the rank-preserving structural failure time model (RPSFTM), as this was expected to give a less biased estimate in the presence of crossover than the unadjusted data.[53] A SA for OS was conducted that included observational studies as well as the RCTs from the primary analysis, which enabled comparison with axitinib. Unfortunately, no RCT or observational studies were identified that reported OS data on sunitinib and so it has not been possible to provide an estimate of its effect on OS. An additional four studies,[43,59,62,63] which included the prior sunitinib subgroup from the AXIS RCT, were suitable for inclusion in the SA for OS.[43] Network diagrams for the primary analysis and SA for OS are presented in *Figures 4* and *Figure 5*, respectively.

The results of the MTC primary analysis for OS did not show statistically significant benefits of any treatment over BSC, but all point estimates were in favour of the active treatment (*Table 8*). Cabozantinib and nivolumab led to longer OS than everolimus (HR 0.66, 95% CrI 0.53 to 0.82; and HR 0.73, 95% CrI 0.60 to 0.89; respectively, see *Table 8*); however, the difference between nivolumab and cabozantinib was not statistically significant (HR 1.12, 95% CrI 0.82 to 1.49). Cabozantinib was associated with the highest probability of being the most effective treatment for prolonging OS:

- cabozantinib, 72.31%
- nivolumab, 24.26%
- everolimus and BSC, 3.43%.

The results of the SA for OS, including data to compare treatments with axitinib, were in keeping with those of the primary analysis (i.e. no statistically significant benefits of treatments over BSC, nivolumab and cabozantinib benefit over everolimus; see *Appendix 4*). Everolimus, cabozantinib and nivolumab all showed longer OS than axitinib (HR 0.74, 95% CrI 0.56 to 0.99; HR 0.48, 95% CrI 0.34 to 0.71; and HR 0.54, 95% CrI 0.38 to 0.77, respectively).

The residual deviance was similar to the number of unconstrained data points in the MTC primary analysis for OS (3 points vs. 3 points, respectively; see *Appendix 5*) suggesting a good model fit. However, the residual deviance was considerably higher than the number of unconstrained data points in the SA (13 points vs. 7 points, respectively). These findings suggest that the results of the SA for OS should be interpreted with caution owing to the poor fit of the MTC model. There was one loop of three studies in the SA: everolimus, axitinib and sorafenib. Investigation of potential inconsistency in the data loop present in the MTC SA for OS suggested that the results were statistically inconsistent ($p > 0.05$; see *Appendix 5*).

TABLE 8 The OS primary analysis. HRs and associated CrIs; HR < 1 favours treatment in the left-hand column

	Treatment			
Treatment	**BSC**	**Nivolumab**	**Cabozantinib**	**Everolimus**
Everolimus	0.53 (0.22 to 1.64)	1.36 (1.12 to 1.67)	1.51 (1.21 to 1.89)	–
Cabozantinib	0.34 (0.14 to 1.12)	0.89 (0.67 to 1.22)	–	0.66 (0.53 to 0.82)
Nivolumab	0.38 (0.16 to 1.23)	–	1.12 (0.82 to 1.49)	0.73 (0.60 to 0.89)
BSC	–	2.62 (0.82 to 6.43)	2.90 (0.89 to 7.19)	1.90 (0.61 to 4.59)

Cells highlighted in green indicate statistically significant results.

Response rate

The MTC for objective response rate (ORR) comprised three RCTs[53,54,57] and allowed only the comparison of cabozantinib, everolimus, nivolumab and BSC (see *Figure 4*). Response data reported in AXIS could not be connected to the network and so have been reported narratively. Response rate was also reported in two observational studies but SAs including non-RCTs were not planned for this outcome.[55,56] The active treatments all showed statistically significant improvements in ORR compared with BSC (*Table 9*). In addition, cabozantinib and nivolumab resulted in statistically significant improvements in ORR compared with everolimus (OR 6.67, 95% CrI 3.28 to 12.78; and OR 6.18, 95% CrI 3.75 to 9.84, respectively). There was no statistically significant difference in ORR between nivolumab and cabozantinib (OR 1.05, 95% CrI 0.41 to 2.18). It should be noted that CheckMate 025 was rated as being at a high risk of bias as a result of the absence of blinding of outcome assessors for response and METEOR was rated as being at an unclear risk of bias for missing data, but evidence for this outcome was otherwise rated as being at a low risk of bias.[54,57] The impact of these potential biases on the overall direction of treatment effects is unknown. Data from AXIS that could not be connected to the MTC showed higher ORR in axitinib-treated (23%) than sunitinib-treated patients (12%), which was similar when ORR was reviewed independently (19% vs. 9%).[43,66]

There were two response rate categories for which it was deemed there were sufficient data for clinically meaningful analyses: stable disease and progressive disease. These analyses used the same three studies as the primary analysis for ORR.[53,54,57] Consistent with the results for ORR, all treatments lowered the odds of having progressive disease and increased the odds of having stable disease, compared with BSC. Cabozantinib significantly lowered the odds of having progressive disease compared with everolimus (OR 0.39, 95% CrI 0.25 to 0.58) and nivolumab (OR 0.27, 95% CrI 0.16 to 0.45); cabozantinib also improved the odds of stable disease compared with nivolumab (OR 2.70, 95% CrI 1.79 to 4.17), but not everolimus (OR 1.18, 95% CrI 0.85 to 1.61). Everolimus lowered the odds of progressive disease compared with nivolumab (OR 0.70, 95% CrI 0.53 to 0.96). All data are shown in *Table 10*. Data from AXIS that could not be connected in the MTC showed similar rates of independently reviewed progressive disease (22% axitinib vs. 21% sunitinib) and higher rates of prolonged stable disease (≥ 20 weeks) with axitinib (27%) than sunitinib (21%).[43,66] Rates of stable disease for < 20 weeks assessed by the investigator were in favour of sunitinib (23% axitinib and 33% sunitinib).

Health-related quality of life

All four of the included RCTs[43,53,54,57] reported data on HRQoL, but the data were not suitable for combining in a MTC. Studies varied in the measures used and analyses undertaken, and they did not create a connected network. The number of available data decreased significantly over the course of the studies owing to progression and, to a lesser extent, questionnaire completion rates. All four studies projected change in HRQoL using repeated measures mixed-effects models to impute data for participants who were not available at each time point. The AXIS, CheckMate 025 and RECORD-1 studies also conducted alternative analyses using pattern mixture models to explore the possibility that the number of, and reasons for, missing data were related to patients' health state (i.e. not missing at random).[43,54,64]

TABLE 9 Objective response rate primary analysis. OR and associated CrI; OR > 1 favours treatment in left-hand column

	Treatment			
Treatment	**BSC**	**Nivolumab**	**Cabozantinib**	**Everolimus**
Everolimus	7.14 (1.32 to 8,216)	0.16 (0.10 to 0.27)	0.15 (0.08 to 0.31)	–
Cabozantinib	42.12 (7.55 to 51,921)	0.95 (0.46 to 2.45	–	6.67 (3.28 to 12.78)
Nivolumab	41.67 (7.56 to 50,276)	–	1.05 (0.41 to 2.18)	6.18 (3.75 to 9.84)
BSC	–	0.02 (0.00002 to 0.13)	0.02 (0.00002 to 0.13)	0.14 (0.0001 to 0.76)

Cells highlighted in green indicate statistically significant results.

TABLE 10 Odds of having stable and progressive disease (for stable disease, OR > 1 favours first treatment; for progressive disease, OR < 1 favours first treatment)

Comparison	OR (95% CrI)	
	Stable disease, OR > 1 favours first treatment	Progressive disease, OR < 1 favours first treatment
Cabozantinib vs. everolimus	1.18 (0.85 to 1.61)	0.39 (0.25 to 0.58)
Everolimus vs. nivolumab	2.33 (1.79 to 3.13)	0.70 (0.53 to 0.96)
Everolimus vs. BSC	4.13 (2.75 to 6.57)	0.28 (0.18 to 0.45)
Cabozantinib vs. BSC	4.76 (2.90 to 8.48)	0.10 (0.06 to 0.20)
Nivolumab vs. BSC	1.73 (1.07 to 3.03)	0.39 (0.23 to 0.69)
Cabozantinib vs. nivolumab	2.70 (1.79 to 4.17)	0.27 (0.16 to 0.45)

Cells highlighted in green indicate statistically significant results.

A summary of HRQoL results from the four included RCTs is given in *Table 11*. HRQoL scores were similar between axitinib and sorafenib in AXIS,[43] regardless of the measure used. Results favoured nivolumab over everolimus in CheckMate 025,[54] a result that was seemingly consistent across the various measures used and type of analysis undertaken. Results in RECORD-1 favoured BSC over everolimus, although this effect was only apparent if models were used to account for data not missing at random.[53] METEOR results favoured everolimus over cabozantinib on a measure of disease-specific quality of life,[57] but scores were similar on two measures of general HRQoL. Completion rates, scales used, analyses undertaken and effect sizes are described in more detail for each study below.

TABLE 11 Summary of HRQoL data from the four included RCTs

	Study, comparison			
	AXIS[43,67]	CheckMate 025[54,68]	METEOR[57,69]	RECORD-1[64,70]
Measure	***Axitinib vs. sorafenib [end-of-treatment MD (95% CI)]***	***Nivolumab vs. everolimus [median change (range) at week 104]***	***Cabozantinib vs. everolimus (change from baseline MD and p-value)***	***Everolimus vs. placebo [time to deterioration: HR (95% CI); results favour placebo]***
FKSI scales	DRS: MD 0.12 (–0.45 to 0.69); $p = 0.68$ FKSI-15: MD 0.35 (–0.63 to 1.34); $p = 0.48$	DRS: nivolumab –2 (–1 to 16) Everolimus 2 (–7 to 15)	FKSI-19: –1.3; $p < 0.0001$	DRS: HR 0.82, 95% CI 0.75 to 0.92; $p = 0.001$
EuroQoL scales	5D index: MD 0.02 (–0.01 to 0.05); $p = 0.19$ VAS: –0.50 (–2.77 to 1.72); $p = 0.65$	–	5D index: MD 0.0; $p = 0.83$ VAS: –0.1; $p = 0.92$	–
EORTC QLQ-C30	–	–	–	Global health status, HR 0.85 (0.75 to 0.96); $p = 0.006$ Physical functioning, HR 0.84 (0.75 to 0.94)'; $p = 0.001$

CI, confidence interval; DRS, Disease Related Symptoms subscale of the FKSI-15; EORTC QLQ-C30, European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire; EQ-5D, EuroQol-5 Dimensions; FKSI, Functional Assessment of Cancer Therapy-Kidney Cancer Symptom Index; MD, mean difference; VAS, visual analogue scale.

Across studies, the proportion of randomised patients in either group with baseline HRQoL measurements was high, ranging from 86.4% (everolimus group of CheckMate 025[54]) to 95.8% (axitinib group of AXIS[43]). METEOR did not state the proportion of the population with baseline measurements.[57] Available measurements as a proportion of the total randomised populations dropped sharply through the course of the studies, primarily owing to deaths and disease progression [axitinib 4.2% and sorafenib 1.9% in AXIS by 19 months (21 cycles),[43] nivolumab 4.9% and everolimus 2.3% in CheckMate 025[54] at 24 months, and everolimus 19% and BSC 4% in RECORD-1[53] by 8 months]. When reported, scale completion rates, defined as the number of completed scales out of the total patients available at each time point, were fairly good; rates remained > 75% in METEOR over 48 weeks,[57] 71–89% in the nivolumab group and 60–90% in the everolimus group in CheckMate.[54] AXIS also took HRQoL measurements at the point when patients' stopped treatment,[43] which were available for 45.2% of the axitinib group and 52.8% of the sorafenib group.

AXIS reported the mean difference (MD) between axitinib- and sorafenib-treated patients on four measures of HRQoL.[43] Patients' baseline Functional Assessment of Cancer Therapy-Kidney Symptom Index-Disease Related Symptoms (FKSI-DRS) subscale scores were described as comparable to the general population in both the axitinib and sorafenib groups of AXIS.[43] Mean scores while on treatment were not significantly different between axitinib and sorafenib, and neither patient group showed a substantial decline during treatment. None of the results from the repeated measures mixed-effects models showed a statistically significant difference between groups [EuroQoL visual analogue scale (EQ-VAS), $p = 0.645$; EuroQol-5 Dimensions (EQ-5D) index score, $p = 0.19$; Functional Assessment of Cancer Therapy Kidney Cancer Symptom Index (FKSI)-15, $p = 0.483$; and FKSI-DRS, $p = 0.675$]. The quality-of-life scores at the point that patients discontinued treatment also did not differ between groups but were significantly below the baseline means, attributed to disease progression. Results from the SAs using pattern-mixed models were not reported in full but were deemed similar to those of the standard mixed-effects model.[67]

RECORD-1 measured HRQoL using the FKSI-DRS subscale and two subscales of the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire (EORTC QLQ-C30).[53] Pattern-mixed models showed that EORTC QLQ-C30 global health status [HR 0.85, 95% confidence interval (CI) 0.75 to 0.96; $p = 0.006$] and physical functioning (HR 0.84, 95% CI 0.75 to 0.94; $p = 0.001$) deteriorated more quickly in patients taking everolimus than those on BSC.[71] Analyses not fitting these models did not show a difference between groups in rate of decline on these measures[64] and the difference between groups was not significant for the FKSI-DRS regardless of which analysis was used.

CheckMate 025 measured HRQoL using the FKSI-DRS subscale, EQ-VAS and EQ-5D index score.[54] For each measure, results were reported for the difference in mean change from baseline up to week 84 (repeated measures mixed-effects models), proportion of patients with meaningful improvement and median time to improvement. The prespecified end point was the number of people in each group with a meaningful improvement of ≥ 2 points on the FKSI-DRS; other analyses were post hoc. Mixed models comparing change from baseline to week 84 favoured nivolumab over everolimus for the FKSI-DRS (MD 1.7, 95% CI 1.2 to 2.1; $p < 0.0001$), EQ-5D index score (MD 0.04, 95% CI 0.02 to 0.07; $p = 0.0003$) and EQ-VAS (MD 5.7, 95% CI 3.8 to 7.7; $p < 0.0001$). The results were deemed consistent when pattern-mixed models were used instead, but the results are not presented, and FKSI-DRS scores, analysed non-parametrically, showed a similar pattern of nivolumab benefit. In the dichotomous analyses using scale cut-off points, more people in the nivolumab group than the everolimus group had clinically significant improvements (an increase on the scale of at least 2 points) over the course of the study on the FKSI-DRS (55% vs. 37%; $p < 0.001$) and the EQ-VAS (53% vs. 39%; $p = 0.0001$); equivalent results for the EQ-5D did not show a difference between groups.[69] The median time to improvement analyses favoured nivolumab, but some CIs were not estimable.

METEOR, comparing cabozantinib with everolimus, reported the EQ-VAS, EQ-5D index score and FKSI-19.[57] Patients randomised to cabozantinib deteriorated more than the everolimus group on the FKSI-19 (mean change from baseline –3.48 cabozantinib vs. –2.21 everolimus; $p < 0.0001$).[28] The scores taken at end of treatment were around 7 points lower than baseline in each group which was attributed to disease progression. Mean change from baseline was similar between the cabozantinib and everolimus groups on the EQ-VAS (MD –0.051; $p = 0.921$) and the EQ-5D index score (MD –0.002; $p = 0.825$).

Adverse events

It was not feasible to conduct a MTC to compare AEs across studies because reporting was inconsistent and there were insufficient numbers of studies reporting on each individual AE. AEs of ≥ grade 3 CTCAE (the criteria defined in the protocol for this review) were only available in AXIS.[43] CheckMate 025, METEOR, and RECORD-1 reported the number or percentage of patients in either group experiencing grade 3 and 4 events (i.e. not including deaths attributed to AEs).[53,54,57] AXIS, CheckMate 025 and RECORD-1 reported treatment-related AEs observed in at least 10% of patients in either group of the study (or just the active everolimus group in RECORD-1),[43,53,54] and METEOR reported events regardless of whether or not they were considered by the investigator to be related to the study treatment.[57] All studies reported the number or percentage of participants experiencing specific AEs according to CTCAE terminology, but only METEOR and CheckMate 025 reported the total number or percentage of patients in each group who had a grade 3 or 4 CTCAE AE of any kind (see *Table 1*).[54,57] AXIS and RECORD-1 used CTCAE version 3.0 criteria,[43,53,72] and CheckMate 025 and METEOR used version 4.0.[54,57,73] The definitions of grade 3 and 4 AEs based on CTCAE version 3 and version 4 are summarised in *Box 1*. Safety assessments took place every 4 weeks in all studies, and more frequently at the start of treatment.

Risk of detection bias is high for AXIS, CheckMate 025 and METEOR because safety assessments were done by investigators who were aware of treatment assignment.[43,54,57] However, safety was overseen by a data monitoring committee in CheckMate 025 and METEOR,[54,57] which may have reduced the risk in those studies. Studies analysed safety data for all patients who received at least one dose of the study medication (between 97% and 99% of the randomised populations). The studies are also rated as being at a low risk of selective reporting biases with regard to AEs.

Table 12 shows a collated list of the 10 most commonly reported CTCAE grade 3 or 4 AEs within each treatment group for each study. Owing to overlap, this gave a list of 25 AEs across studies and treatments. AEs shown in green indicate a statistically higher percentage of patients in that group who experienced that event than patients in the other treatment group. RECORD-1 reported grade 3 AEs separately from grade 4 AEs, which were summed to make the data more comparable to the other studies.[53] The RECORD-1 grade 4 events constituted none, or a very small proportion of, the summed events (see *Appendix 6*).[53]

Patients who received everolimus in RECORD-1 had higher rates of anaemia, raised cholesterol, hyperglycaemia, infections, lymphopenia, decreased phosphate, pneumonitis and stomatitis compared with those who received BSC (placebo).[53] Compared with cabozantinib in METEOR,[57] those given everolimus also had higher rates of anaemia and hyperglycaemia, but rates of stomatitis were similar between the two treatments. Raised cholesterol, infections, lymphopenia, decreased phosphate and pneumonitis were not reported in METEOR (i.e. not observed in > 10% of either group).[57] Compared with nivolumab in CheckMate 025,[54] the number of patients who experienced any grade 3 or 4 AE was significantly higher in those who received everolimus (36.5%) than those who received nivolumab (18.7%). More patients who

BOX 1 Summary of CTCAE version 3.0 and version 4.0 criteria for grade 3 and 4 AEs

CTCAE version 3.0 grade 3 = severe.

CTCAE version 3.0 grade 4 = life-threatening or disabling.

CTCAE version 4.0 grade 3 = severe or medically significant but not immediately life-threatening *or* hospitalisation *or* prolongation of hospitalisation indicated *or* disabling *or* limiting self-care activities of daily living.

CTCAE version 4.0 grade 4 = life-threatening consequences *or* urgent intervention indicated.

TABLE 12 Common grade 3–4 treatment-related[a] AEs across studies

AE	Study							
	RECORD-1[53]		METEOR[57]		CheckMate 025[54]		AXIS[43]	
	Everolimus, % ($n = 274$)	BSC, % ($n = 137$)	Everolimus, % ($n = 322$)	BSC, % ($n = 331$)	Everolimus, % ($n = 397$)	Nivolumab, % ($n = 406$)	Axitinib, % ($n = 359$)	Sorafenib, % ($n = 355$)
Any AE, grade 3 or 4	–	–	59.9	71.0	36.5	18.7	–	–
Abdominal pain	–	–	1.5	3.6	–	–	–	–
Anaemia	13.1	5.8	16.5	5.7	7.8	1.7	–	–
Anorexia	–	–	–	–	–	–	4.2	2.0
Asthenia	–	–	2.5	4.5	–	–	4.2	2.3
Back pain	–	–	2.2	2.4	–	–	–	–
Cholesterol increased	4.0	0.0	–	–	–	–	–	–
Decreased appetite	–	–	0.9	3	1.0	0.5	4.2	2
Diarrhoea	–	–	2.2	11.5	1.3	1.2	11.1	7.6
Dyspnoea	6.9	2.9	4.3	3	0.7	0.7	–	–
Fatigue	5.1	3.6	7.5	9.1	2.8	2.5	10.3	3.9
Hand–foot syndrome	–	–	0.9	8.2	–	–	5.6	17.2
Hyperglycaemia	15.7	1.5	5.0	0.9	3.8	1.2	–	–
Hypertension	–	–	3.7	14.8	–	–	16.7	12.1
Hypertriglyceridemia	–	–	3.1	1.2	5.0	0	–	–
Hypomagnesaemia	–	–	0	4.8	–	–	–	–
Infections	9.9	1.5	–	–	–	–	–	–
Lymphopenia	17.9	5.1	–	–	–	–	–	–
Mucosal inflammation	–	–	3.4	1.5	3.0	0	1.4	0.8

continued

TABLE 12 Common grade 3–4 treatment-related[a] AEs across studies (*continued*)

AE	Study							
	RECORD-1[53]		METEOR[57]		CheckMate 025[54]		AXIS[43]	
	Everolimus, % ($n = 274$)	BSC, % ($n = 137$)	Everolimus, % ($n = 322$)	BSC, % ($n = 331$)	Everolimus, % ($n = 397$)	Nivolumab, % ($n = 406$)	Axitinib, % ($n = 359$)	Sorafenib, % ($n = 355$)
Nausea	–	–	0.3	4.5	0.8	0.2	1.7	0.8
Phosphate decreased	5.8	0.0	–	–		–	–	–
Pneumonitis	4.0	0.0	–	–	2.8	1.5	–	–
Proteinuria	–	–	0.6	2.4		–	3.1	1.1
Rash	–	–	0.6	0.6	0.5	0.5	0.3	3.7
Stomatitis	4.7	0.0	2.2	2.4	4.3	0	1.4	0.3
Weight decreased	–	–	0	2.7	–	–	3.3	2.5

a METEOR AEs were not necessarily treatment related.
Data are the percentage of patients experiencing a given AE out of the total included in the safety analyses; empty cells show unreported AEs, usually because the AE was not reported frequently enough in the study to be included in the AE table; cells with green data indicate a statistically significant difference between groups. Percentages for RECORD-1 are the sum of grade 3 and 4 events, and AXIS data also include CTCAE grade 5 events.

received everolimus in CheckMate 025 had anaemia, hypoglycaemia and stomatitis as in RECORD-1,[53,54] and there were also higher rates of hypertriglyceridemia and mucosal inflammation. CheckMate 025 did not report an increased rate of cholesterol, infections, lymphopenia or decreased phosphate, but rates of pneumonitis were similar between groups.[54]

More patients who received cabozantinib (71.0%) in METEOR had any grade 3 or 4 AEs than those who received everolimus (59.9%).[57] The higher rates in this study compared with CheckMate 025 may be due to the wider definition of AEs,[54] including AEs that were not thought to be treatment-related, but may equally be indicative of differences in study populations. Patients who received cabozantinib in METEOR had higher rates of diarrhoea, hand–foot syndrome, hypotension, hypomagnesaemia, nausea and weight loss than those who received everolimus.[57]

Of the AEs reported in CheckMate 025,[54] none was reported by significantly more people who received nivolumab than those who received everolimus. Just under 19% of the nivolumab group had any grade 3 or 4 AE, and the most commonly reported grade 3 or 4 AEs were fatigue (2.5%), anaemia (1.7%), pneumonitis (1.5%), hyperglycaemia and diarrhoea (both 1.2%); all other grade 3 and 4 AEs were reported in < 1% of those who received nivolumab.

Subgroup analyses

Two subgroup analyses were planned for the primary outcomes (PFS and OS): prior therapies and baseline prognostic scores (e.g. MSKCC). Formal analysis of these moderators using MTC was limited by what was reported in the studies and whether or not the treatments could be connected in a network. For both outcomes, we were able to perform a MTC to investigate relative treatment effects across MSKCC categories (favourable, intermediate, poor). For prior therapies, we were able perform a MTC to relative effectiveness for people who had received one prior TKI compared with those who had received two or more prior TKIs.

Subgroup data (PFS and OS) for ECOG prognostic score were reported in METEOR;[57] RECORD-1 reported PFS results by prior sunitinib, prior sorafenib and both.[64] These data could not be combined with any other studies and have been described narratively. Subgroup data by prior therapies in AXIS were not used because the sunitinib subgroup was the basis for including the study in this review; the other subgroups (prior cytokines, bevacizumab or temsirolimus) were not relevant to the review question.[43] Subgroup analyses from observational data were not part of the primary analyses and so have not been described, but some were available in Vogelzang *et al.*[61] (prognostic score for PFS and OS) and Porta *et al.*[62] (prognostic score and prior therapy for PFS only).

Baseline prognostic score

The prognostic scores used at baseline varied across the four included RCTs and included ECOG performance status, MSKCC and Karnofsky performance status. RECORD-1 and METEOR were the only studies that reported subgroup data on PFS by baseline MSKCC prognostic score.[53,57] The MSKCC groups with data were defined as favourable, intermediate and poor prognosis. The results of the MTC analyses for PFS demonstrate consistent treatment benefit with both cabozantinib and everolimus compared with BSC across all three MSKCC categories (favourable, intermediate and poor). PFS subgroup data in METEOR that could not be combined in a MTC showed a larger effect in favour of cabozantinib than everolimus for people with an ECOG score of 0 at baseline (HR 0.46, 95% CI 0.36 to 0.59) than those with a score of 1 (HR 0.64, 95% CI 0.46 to 0.90).[57]

Overall survival data for the baseline prognostic score MTC subgroup analysis by MSKCC risk were from two studies (CheckMate 025 and METEOR).[54,57] These results suggested a trend in favour of cabozantinib and nivolumab over everolimus irrespective of MSKCC baseline score, although the results failed to reach statistical significance in some subgroups. However, this is to be expected because the individual studies in the MTC were not powered sufficiently for these subgroup analyses. OS subgroup data in METEOR that could not be combined in a MTC showed a slightly larger effect in favour of cabozantinib over everolimus

for people with an ECOG score of 0 at baseline (HR 0.65, 95% CI 0.49 to 0.87) than those with a score of 1 (HR 0.72, 95% CI 0.51 to 1.02), but CIs were overlapping.[57]

Prior therapy

The only data available for analyses by number of prior therapies were data on patients with one prior TKI and data on people with more than two prior TKIs. Two studies (METEOR and RECORD-1) were included in the MTC of PFS. The MTC consistently demonstrated better PFS with cabozantinib than with everolimus and either of the active treatments compared with BSC, irrespective of number of prior therapies (see *Appendix 7*).[57,74] OS subgroup data in RECORD-1 that could not be combined in a MTC showed very similar effects of everolimus over placebo (BSC) for people previously treated with sorafenib only, sunitinib only, or both (HR 0.29, 0.30 and 0.28, respectively, CIs not given).[64]

There were only two studies (CheckMate 025 and METEOR) reporting suitable data for analysis of OS by number of prior TKI therapies.[54,57] The MTC results for OS indicate no statistically significant difference in efficacy between cabozantinib and nivolumab irrespective of the number of prior TKI therapies (see *Appendix 7*). Treatment with cabozantinib and nivolumab both resulted in a longer OS than with everolimus, but the differences were only statistically significant after one prior TKI.

Summary of the results of the review of clinical effectiveness

In a MTC for the primary analysis of PFS, cabozantinib was associated with a statistically significant improvement in PFS compared with everolimus with a HR of 0.51 (95% CrI 0.41 to 0.63). Cabozantinib and everolimus both showed statistically significant benefits over BSC (HR 0.17, 95% CrI 0.12 to 0.24; and HR 0.33, 95% CrI 0.25 to 0.43, respectively). It was not possible to include nivolumab in the analyses of PFS because PHs did not hold for PFS in CheckMate 025,[54] which was the only study evaluating nivolumab included in this review. A SA connected axitinib and sunitinib to the network by including five reasonable-quality non-RCTs[5,9–12] and, thus, a third RCT.[2] This analysis showed statistically significant benefits for PFS for all active treatments over BSC (everolimus HR 0.33, 95% CrI 0.25 to 0.43; cabozantinib HR 0.17, 95% CrI 0.12 to 0.24; axitinib HR 0.31, 95% CrI 0.21 to 0.44; and sunitinib HR 0.27, 95% CrI 0.17 to 0.40). Cabozantinib showed a statistically significant benefit for PFS against all other treatments: everolimus (HR 0.51, 95% CrI 0.41 to 0.63), sunitinib (HR 0.63, 95% CrI 0.44 to 0.95), axitinib (HR 0.54, 95% CrI 0.40 to 0.76) and BSC (HR 0.17, 95% CrI 0.12 to 0.24). None of the differences in PFS between sunitinib, everolimus and axitinib was statistically significant. A second SA, including two additional non-RCTs rated as being at a critical risk of bias, was consistent with the primary analysis and first SA. Cabozantinib was found to have a 99% probability of being the most effective treatment for improving PFS.[7,8]

The results of the MTC for OS suggest that both cabozantinib and nivolumab significantly prolong OS compared with everolimus and that there is no statistically significant difference between nivolumab and cabozantinib:

- cabozantinib versus everolimus HR 0.66 (95% CrI 0.53 to 0.82)
- nivolumab versus everolimus HR 0.73 (95% CrI 0.60 to 0.89)
- nivolumab versus cabozantinib HR 1.12 (95% CrI 0.82 to 1.49).

The results of the SA for OS, including data to compare treatments with axitinib, were in keeping with those of the primary analysis (no statistically significant benefits of treatments over BSC; nivolumab and cabozantinib benefit over everolimus). Everolimus, cabozantinib and nivolumab all showed longer OS compared with axitinib (HR 0.74, 95% CrI 0.56 to 0.99; HR 0.48, 95% CrI 0.34 to 0.71; and HR 0.54, 95% CrI 0.38 to 0.77, respectively). Data were not available to provide an OS estimate for sunitinib compared with the other treatments. However, it should be noted that there was a statistically significant inconsistency in the network for this analysis, which is discussed further in *Chapter 4*.

Analyses of response rate confirmed the clinical effectiveness of cabozantinib, everolimus and nivolumab compared with BSC (placebo) and suggests that cabozantinib and nivolumab were the most effective treatments. Cabozantinib and nivolumab resulted in statistically significant improvements in ORR compared with everolimus (OR 6.67, 95% CrI 3.28 to 12.78; and OR 6.18, 95% CrI 3.75 to 9.84, respectively), and the difference between the ORR for the two treatments was not statistically significant (OR 1.05, 95% CrI 0.41 to 2.18). The active treatments all resulted in statistically significant improvements in ORR compared with BSC, but small numbers of events led to very wide CrIs:

- everolimus versus BSC (OR 7.14, 95% CrI 1.32 to 8216)
- cabozantinib versus BSC (OR 42.12, 95% CrI 7.55 to 51,921)
- nivolumab versus BSC (OR 41.67, 95% CrI 7.56 to 50,276).

Data on HRQoL and AEs were not suitable for combining in MTC analyses.

In summary, the results of the analyses of PFS, OS and ORR suggest that cabozantinib is the most effective treatment, closely followed by nivolumab, with little difference between axitinib, everolimus and sunitinib. All of the active treatments appear to be more effective than BSC.

Chapter 4 Assessment of cost-effectiveness

Assessment of the cost-effectiveness of second-line treatments for amRCC was undertaken through carrying out a systematic review of the published research literature and previous NICE TAs (see *Systematic review of existing cost-effectiveness evidence*) and through development of a de novo economic analysis (see *Independent economic assessment*).

Systematic review of existing cost-effectiveness evidence

This section provides a review of the existing cost-effectiveness evidence, both published and presented within previous NICE TAs, for second-line treatments of amRCC covered in the scope of this HTA report.

- The systematic review of published cost-effectiveness evidence carried out by the assessment group (AG), together with the search results, is presented in the section *Systematic review of published cost-effectiveness evidence*
- The identified cost-effectiveness studies and a description of the studies is presented in the section *Overview of the identified cost-effectiveness studies*
- The cost-effectiveness evidence presented in previous NICE TAs on second-line treatment of RCC is summarised in the section *Review of cost-effectiveness evidence in previous National Institute for Health and Care Excellence technology appraisals on second-line treatment of renal cell carcinoma.*

Systematic review of published cost-effectiveness evidence

A systematic literature review was carried out in February 2016 to identify full economic evaluations and costing studies relevant to the decision problem. The following electronic databases were searched:

- MEDLINE (via Ovid)
- EMBASE (via Ovid)
- DARE (via DARE)
- NHS Economic Evaluation Database (via NHS EED).

Databases were searched from inception to identify all evidence related to the relevant interventions. The search strategy combined terms capturing the interventions or comparators of interest (axitinib, everolimus, nivolumab, sorafenib and sunitinib) and the target condition (RCC). Search terms such as cost-effectiveness, cost–utility, cost and health state utility values (HSUVs) were applied to capture the study designs of interest. No language, setting or country restrictions were applied to the search strategy. The search strategies and results are reported in *Appendix 1*.

In addition to searching the aforementioned databases, the following sources of potentially relevant publications were also explored.

- The NICE TA website was searched for any published TA for second-line treatment of RCC that had not already been identified via the database searches or that could potentially include additional HRQoL data. Results of the search are reported in *Review of cost-effectiveness evidence in previous National Institute for Health and Care Excellence technology appraisals on second-line treatment of renal cell carcinoma*.
- Reference lists of key identified studies were reviewed for any potentially relevant studies.

All titles and abstracts of the papers identified through the searches outlined above were independently assessed for inclusion by two health economists using the criteria presented in *Appendix 2*.

Sources of evidence for different interventions considered relevant in second-line RCC were identified in the search, to be used as alternative resources in case the relevant evidence base identified through the search was not sufficient. These were collected for bevacizumab, interferon-alpha, pazopanib, temsirolimus and tivozanib, as specified in the protocol.[42]

Search results

The systematic review identified 633 potentially relevant citations after deduplication was carried out. The titles and abstracts of the papers were reviewed independently by two reviewers (FS and PC); conflicts were resolved by discussion between the reviewers and, when necessary, by a third reviewer (MB). A total of 464 papers were excluded based on title and abstract.

A total of 169 full-text papers were considered potentially relevant and were reviewed independently by two reviewers. Conflicts were resolved by third-party arbitration with a third reviewer. After review of the studies, 11 studies were included. Out of the 11 included citations, 10 were economic evaluations and one was a HRQoL paper. No UK costing studies were identified. The remaining 158 studies were excluded for the following reasons:

- abstract with insufficient methodological details and no additional details published elsewhere ($n = 104$)
- non-UK costing study ($n = 15$)
- wrong population ($n = 13$)
- no relevant outcomes reported ($n = 6$)
- irretrievable ($n = 6$)
- study not available in English ($n = 4$)
- duplicate ($n = 5$)
- wrong study design (i.e. not economic evaluation, costing study or quality-of-life study) ($n = 3$)
- letter/commentary ($n = 1$)
- systematic review ($n = 1$).

Non-UK costing studies were excluded given that the review of cost-effectiveness evidence presented in previous NICE TA on second-line treatment of RCC was deemed sufficient to capture the resource use in advanced RCC in the UK.

The PRISMA flow diagram for the search is reported in *Figure 6*.[75] The full references of the studies excluded after review of full papers are reported in *Appendix 10*.

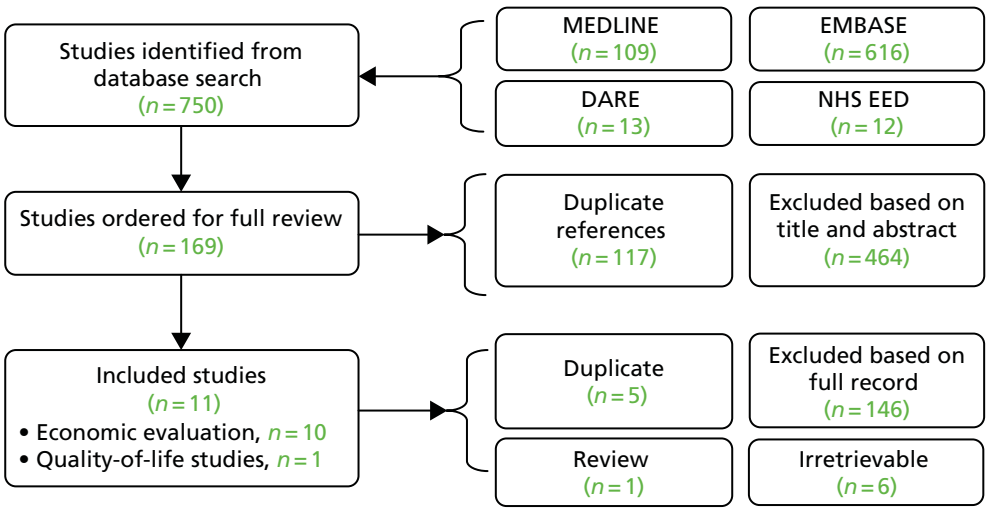


FIGURE 6 The PRISMA flow diagram.

Overview of the identified cost-effectiveness studies

A total of 10 studies that reported economic evaluations were identified from the systematic review. The studies are described in further detail in *Narrative description of published cost-effectiveness studies* and the complete data extraction tables are presented in *Appendix 8*. Nine studies reported cost–utility analyses with cost per quality-adjusted life-year (QALY) as the key outcome,[76–84] while one study was a budget impact analysis.[85] The studies assessed the cost-effectiveness of various second-line treatment options in patients with RCC. Some of the studies included a description of previous therapies received by patients. In Paz-Ares *et al.*[83] and Purmonen *et al.*,[84] patients in the model were assumed to have failed previous cytokine therapy (i.e. interleukin 2 or interferon-alpha first line).[83,84] The target populations in the studies by Casciano *et al.*[80] and Lopes *et al.*[85] were patients refractory to sunitinib while, in the study by Mihajlovic *et al.*,[79] it was patients who had progressed after receiving sunitinib or sorafenib. In the study by Petrou,[76] patients were assumed to have previously received sunitinib or cytokines.

None of the studies included all of the comparators originally specified in the protocol. Six studies included sorafenib,[76–78,80–82] two included sunitinib,[83,84] three studies included everolimus[79,80,85] and only one study had axitinib as a comparator.[76] None of the studies identified by the search included nivolumab as a treatment.

Markov models were used in all of the cost–utility analyses to estimate cost-effectiveness. A critique of these analyses using the NICE reference checklist and the Philip's checklist is presented in *Appendix 9*.[86,87] The health states included across the models to represent the disease pathway were similar; the model structures generally included a health state representing PFS or stable disease, a health state representing disease progression, and an absorbing state for death. In the study by Casciano *et al.*,[80] stable disease was further divided into stable disease with and stable disease without AEs in order to model costs associated with AEs. Two of the studies were carried out in a UK setting: Hoyle *et al.*[81] summarised in the next subsection and Thompson Coon *et al.*[82] The latter is described in *Review of cost-effectiveness evidence in previous National Institute for Health and Care Excellence technology appraisals on second-line treatment of renal cell carcinoma* along with other published NICE TAs.[26,28,29,88–90]

Narrative description of published cost-effectiveness studies

Petrou[76]

This study assessed the cost-effectiveness of axitinib (5 mg, twice a day) compared with sorafenib (400 mg, twice a day) in patients with RCC who have been previously treated with sunitinib or cytokines. The analysis was carried out from a Cypriot health-care payer perspective. Costs and outcomes were discounted at a 3.5% annual discount rate. The cost year was not reported.

A Markov model with monthly cycles was used to estimate cost-effectiveness of axitinib compared with sorafenib over a 10-year time horizon. The model included the health states PFS, progressed disease (PD) and death.

Treatment effectiveness was based on survival data from the AXIS trial, which assessed the effectiveness of axitinib compared with sorafenib in a total of 723 patients (361 in the axitinib arm and 262 in the sorafenib arm) for a follow-up period of 3 years.[66] The authors reported that it was the only Phase III RCT with axitinib as a comparator and > 100 patients per arm that they identified at the time.

The HSUVs were estimated based on EQ-5D data from the AXIS trial and were 0.69 and 0.61 for PFS and PD, respectively.[67] The model included pharmaceutical and medical costs.

The incremental cost-effectiveness ratio (ICER) for axitinib compared with sorafenib was €87,936 per QALY gained. Deterministic and probabilistic sensitivity analyses (PSAs) were carried out. The one-way sensitivity analysis (OWSA) indicated that the ICER is sensitive to the price of axitinib and utility values and to a lesser extent to PFS and OS estimates. The probability of axitinib being cost-effective at a willingness-to-pay (WTP) threshold of €60,000 per QALY was 13%.

Petrou and Talias[77]

This study assessed the cost-effectiveness of sorafenib (400 mg, twice daily) compared with BSC for treatment of second-line RCC. The analysis was carried out from a Cypriot health-care payer perspective. Costs and outcomes were discounted at an annual rate of 3.5%. Costs were reported in 2012 prices.

A Markov model with monthly cycles was used to carry out the analysis over a time horizon of 10 years. The model included the health states PFS, PD and death.

Effectiveness data were obtained from TARGET (Treatment Approaches in Renal Cancer Global Evaluation Trial),[91] a multinational, multicentre Phase III RCT assessing the effectiveness of sorafenib compared with BSC in 903 patients, with a follow-up period of 18 months. Petrou and Talias[77] reported that this was the only trial identified which compared sorafenib with BSC. The median age of patients in the trial was 58 years. Nearly all (99%) of the patients had clear-cell carcinoma and 93% of them had nephrectomy prior to trial entry.[91] The duration of PFS and disease progression in the trial was used to estimate the transition probabilities in the economic model.

The HSUVs used in the model were the same as those reported by Thompson Coon *et al.*[82] in the NICE TA report (TA178), which were valued using UK EQ-5D tariffs. The values for PFS and PD were 0.76 [standard error (SE) 0.03] and 0.68 (SE 0.04), respectively.[29] The model included medical and drug costs.

The probabilistic ICER for sorafenib and BSC compared with BSC was €102,059 per QALY gained. OWSA was carried out. The OWSA showed the ICER to be sensitive to the price of sorafenib, utility estimates and PFS estimates.

Petrou and Talias[78]

This study assessed the cost-effectiveness of sorafenib (400 mg, twice daily) compared with BSC for the treatment of second-line RCC. The analysis was carried out from a Cypriot health-care payer perspective. Costs and outcomes were discounted at an annual rate of 3.5%. Costs were reported in 2012 prices.

A Markov model with monthly cycles was used to carry out the analysis over a time horizon of 10 years. The model included the health states PFS, PD and death.

Effectiveness data were obtained from TARGET, a Phase III RCT assessing the effectiveness of sorafenib compared with BSC.[91] The duration of PFS and disease progression in the trial was used to estimate the transition probabilities in the model.

The HSUVs used in the model were the on EQ-5D data reported by Motzer *et al.*[92] and the authors reported that UK EQ-5D tariffs were used for valuation. The values for PFS and PD were 0.76 (SE 0.03) and 0.68 (SE 0.04), respectively.[92] The model included medical and drug costs.

The probabilistic ICER for sorafenib compared with BSC was €102,616 per QALY. OWSA was carried out and showed the ICER to be sensitive to drug costs and OS estimates.

Mihajlovic et al.[79]

The study by Mihajlovic *et al.*[79] assessed the cost-effectiveness of everolimus (10 mg, once daily) and BSC compared with BSC for the treatment of second-line RCC. The analysis was carried out from a Serbian health-care payer perspective. Costs and outcomes were discounted at an annual rate of 3% and 1.5%, respectively. Costs were reported in 2013 prices.

A Markov model with 8-week long cycles was used to carry out the analysis, over a total of 18 cycles. The health states included in the model were stable disease, PD and death.

Effectiveness was estimated based on survival data reported in the RECORD-1 trial, a Phase III RCT assessing the effectiveness of everolimus compared with BSC in 416 patients. A log-normal distribution was fitted to PFS trial data and a Weibull distribution was used to model OS data from the RECORD-1 trial in order to estimate time-dependent transition probabilities to be used in the economic model.[93]

The authors reported that they used utility values reported in NICE TA178.[82] However, the actual values used were not reported in the paper. The costs included in the model were drug acquisition costs, management costs and AE costs.

The probabilistic ICER for everolimus compared with BSC was €86,978 per QALY gained. Deterministic and PSAs were carried out. The OS HR was found to be the most influential parameter in the model. The results were also sensitive to choice of distribution for fitting PFS and OS data, and to the uncertainty surrounding the PFS and OS estimates.

Lopes et al.[85]

Lopes *et al.*[85] carried out a budget impact analysis to assess the impact of introducing everolimus (10 mg/day) to treat patients with advanced RCC who failed to respond to, or have become intolerant to, sunitinib or sorafenib. The analysis was carried out from a US payer perspective and considered only pharmacotherapeutic treatments for RCC (i.e. systemic chemotherapy and targeted treatments). Existing treatments considered in the analysis to reflect the current market situation were bevacizumab, interferon, interleukin, pazopanib, sorafenib, sunitinib and temsirolimus. A Microsoft Excel® (Microsoft Corporation, Redmond, WA, USA)-based cross-sectional budget impact model was used. Costs were estimated for the periods of April 2008 to March 2009, and October 2009 to September 2010 to reflect the periods before and after expected uptake of everolimus, respectively, thus allowing a comparative analysis of the costs of everolimus. Costs were reported in 2010 USD.

A hypothetical cohort of 1 million patients was assumed to receive everolimus in the model. This assumption was based on real-time drug utilisation data from 36 states across the USA.

Costs included in the model were drug costs, administration and AE management costs. Costs of BSC and palliative care were not included in the model.

The results of the analyses showed that in the market following the introduction of everolimus, the total cost decreased from US$7,050,158 to US$6,741,642, yielding savings of US$308,516 (compared with a market in which everolimus is not available). The budget impact per member per month cost was US$0.03 and the budget impact per member per year cost was US$0.31. Therefore, the authors concluded that introduction of everolimus had a minimal impact on the budget. Sensitivity and scenario analyses were carried out and indicated the estimated budget impact to be relatively stable.

Casciano et al.[80]

The study by Casciano *et al.*[80] assessed the cost-effectiveness of everolimus (10 mg per day) compared with sorafenib (800 mg per day). The analysis was carried out from a US payer perspective. Costs and outcomes were discounted at an annual rate of 3% and costs were reported in 2010 USD.

A Markov model with 8-week cycles was used to estimate the cost-effectiveness of everolimus compared with sorafenib over a time horizon of 6 years. The health states included in the model were stable disease with no AEs, stable disease with AEs, PD and death.

Transition probabilities were estimated using a subset of the RECORD-1 trial patient population receiving everolimus after sunitinib, and a comparable population receiving sorafenib in a single-arm Phase II study.[64,94]

The HSUVs for stable disease (no AEs) and for PD were obtained from the NICE TA178[29] report while those for stable disease with AEs were obtained from a study of advanced small lung cancer.[95] The HSUVs for stable disease with no AEs, stable disease with AEs, PD and death were 0.76 [standard deviation (SD) 0.03], 0.71 (SD 0.04), 0.68 (SD 0.04) and 0, respectively.[95,96]

The costs of monitoring, blood tests, CT scans, AEs and analgesics were included in the model. Post-discontinuation treatments were also included in the model.

A deterministic ICER of US$89,160 per QALY and a probabilistic ICER of US$76,496 per QALY gained were estimated. The authors reported carrying out OWSAs on key parameters of the deterministic model as well as a PSA. The OWSAs identified post-discontinuation treatments as one of the main drivers in the model. Assumptions surrounding treatment dose intensity and mortality rate also had an impact on the results.

Hoyle et al.[81]

Hoyle *et al.*[81] reports the methods and results of an economic analysis that compared sorafenib (400 mg, twice a day) with BSC for the second-line treatment of RCC. The analysis was carried out from a UK NHS and Personal Social Services (PSS) perspective, over a time horizon of 10 years. Costs and outcomes were discounted at an annual rate of 3.5% and costs were inflated to 2007/8 prices. A Markov model with 6-week cycles was used to estimate the cost-effectiveness of sorafenib compared with BSC, which included PFS, PD and death health states.

Effectiveness data were obtained from TARGET which is a multinational, multicentre Phase III RCT assessing the effectiveness of sorafenib compared with BSC in 903 patients.[91] Weibull curves were fitted to PFS and OS KM data from the trial. HRs were used in the model as measures of relative effectiveness.

Costs considered in the analysis included drug costs, consultant visits, general practitioner (GP) visits, blood tests, CT scans, community nurse visits and pain medication. As the trial did not report treatment compliance, the authors assumed that patients received 100% of sorafenib doses and that the first pack was free, which is in line with an agreement between the pharmaceutical company and the Department of Health. Total costs for patients treated with sorafenib and patients on BSC were estimated to be £23,860 and £3797, respectively.

The HSUVs used in the model were obtained from a Phase II trial of sunitinib,[92] using EQ-5D tariffs as reported in the NICE TA178 report.[82] The HSUVs associated with PFS and PD were 0.76 and 0.68, respectively.[82]

The ICER for sorafenib compared with BSC was estimated to be £75,398. The authors carried out OWSAs and a PSA. According to the OWSAs, the cost-effectiveness results were sensitive to changing assumptions related to fitting of OS and PFS curves, in addition to HSUVs and drug costs. The results of the PSA indicated that the probability of sorafenib being cost-effective at a WTP threshold of £30,000 per QALY was 0%.

Paz-Ares et al.[83]

The study by Paz Ares *et al.*[83] assessed the cost-effectiveness of sunitinib (50 mg daily for 4 weeks followed by 2 weeks of rest) compared with BSC as a second-line treatment of RCC in patients who are intolerant or have not responded to cytokines. BSC was reported to comprise analgesics and megestrol acetate (320 mg per day). The analysis was carried out from a Spanish NHS perspective, over a time horizon of 10 years. Costs and outcomes were discounted at an annual rate of 3.5% and costs were reported in 2007 prices. A Markov model was used to estimate the cost-effectiveness of sunitinib compared with BSC and included the following health states: PFS, PD and death.

Survival time and time to first progression for patients assumed to receive sunitinib were taken from an open-label single-arm Phase II sunitinib trial in patients refractory to cytokines reported by Motzer *et al.*[92] The authors stated that time to second progression was also taken from the same trial and assumed to be equal for sunitinib and BSC. However, 'second progression' was not defined and it is unclear what it refers to.

As the authors could not find a trial comparing sunitinib with placebo in addition to BSC in this population, a retrospective analysis of data from the American Surveillance, Epidemiology and End Results (SEER) programme of the National Cancer Institute Database and from the Medicare database was carried out to estimate the effectiveness of BSC.[97,98] These databases monitored sunitinib-treated patients with advanced RCC who had experienced disease progression on cytokines treatment and who had been diagnosed between 1997 and 2002. Patients were followed until death.

Health state utility values were derived from the Phase II, single-arm sunitinib trial.[92] The same utility values for respective health states were assumed for sunitinib and BSC (i.e. there was no direct impact assumed for being on active treatment compared with being on BSC). However, utility values were assumed to differ during and after progression and patients in the sunitinib group were assumed to experience a lower utility compared with BSC patients when experiencing AEs. The values used were 0.764 and 0.731 for patients who were alive without progression and those whose disease had progressed, respectively. Patients on sunitinib experiencing AEs were assumed to have a 5.7% reduction in utility. The ICER for sunitinib compared with BSC was €34,196 per QALY.

The authors reported carrying out OWSAs and a PSA. The OWSAs indicated that the cost-effectiveness results were sensitive to assumptions surrounding drug costs, survival estimates and the utility values used for sunitinib. The PSA showed that, at a WTP threshold of €50,000 per QALY gained, the probability of sunitinib being cost-effective compared with BSC was 99%.

Purmonen et al.[84]

This study assessed the cost-effectiveness of sunitinib (50 mg daily for 4 weeks followed by 2 weeks of rest) compared with BSC (including palliative chemotherapy) for the second-line treatment of RCC in patients who were refractory to cytokines. The analysis was carried out from a Finnish health-care payer perspective. Costs and QALYs were discounted at an annual rate of 5% and costs were reported in 2005 prices.

The analysis was carried out using a Markov model with monthly cycles over a lifetime horizon (5 years). The health states included in the model were no new progression-related events, history of progression-related events and death.

Efficacy of sunitinib was estimated based on survival data from two single-arm Phase II trials.[92,99] The trials reported PFS and, therefore, the data were pooled to estimate PFS in the model for sunitinib. Median OS was reported by one trial and these data were used to inform OS estimates in the economic model.[92] PFS was based on data from a total of 168 patients and OS from 63 patients. Weibull curves were fitted to PFS and OS KM data from the trials. Effectiveness and resource use related to BSC were estimated using data from the medical records of 39 patients from two Finnish university hospitals.

The authors reported that the HSUVs used in the model were derived from EQ-5D data collected in the Phase II sunitinib trial. The HSUVs used were 0.764 and 0.731 for before new progression and after progression, respectively. Medication, examinations and hospital unit costs were included in the economic model.

The ICER for sunitinib compared with BSC was €43,698 per QALY gained. Deterministic OWSAs and PSAs were carried out and the authors reported that assuming treatment continuation for a month after progression had the greatest impact on the results, increasing the ICER to €49,000 per QALY. The probability of sunitinib being cost-effective at a WTP of €45,000 per QALY gained was around 70% in a 5-year time horizon.

Review of cost-effectiveness evidence in previous National Institute for Health and Care Excellence technology appraisals on second-line treatment of renal cell carcinoma

This section presents the review of previous NICE TAs on second-line treatments for an amRCC, identified from the NICE website (www.nice.org.uk/guidance/published). In the February 2016 systematic review carried out for this report, three relevant appraisals were identified, and a further two appraisals that were published subsequently were also included. The five included appraisals were as follows: a MTA, which included two relevant interventions, and four STAs. The MTA evaluated sorafenib and sunitinib as monotherapies for the second-line treatment of amRCC (TA178) as well as other first-line treatments options that are not within the scope of this review.[29] The four STAs evaluated everolimus for the second-line treatment of amRCC (TA219), axitinib for the second-line treatment of amRCC (TA333), nivolumab for the second-line treatment of amRCC (TA417) and cabozantinib for the second-line and third-line treatment of amRCC (TA463).[28,88–90]

The economic evidence presented within these appraisals was considered to be a relevant source of information, and the cost-effectiveness analyses presented within them are summarised in sections *Multiple technology appraisal number 178: bevacizumab (first line), sorafenib (first and second line), sunitinib (second line) and temsirolimus (first line) for treatment of advanced and/or metastatic renal cell carcinoma* to *Single technology appraisal TA463: cabozantinib for treating advanced renal cell carcinoma in adults who have received at least one prior vascular endothelial growth factor-targeted therapy*.

Multiple technology appraisal number 178: bevacizumab (first line), sorafenib (first and second line), sunitinib (second line) and temsirolimus (first line) for treatment of advanced and/or metastatic renal cell carcinoma

For the purposes of this HTA report, the review of TA178 will only focus on the areas relating to second-line treatments of RCC; that is, the evaluation of sorafenib (second line only) and sunitinib. We begin by presenting a brief description of the company's submission on the different drugs, followed by the AG analysis.

Company's submission: sorafenib (Bayer, Leverkusen, Germany)

The holder of the marketing authorisation for sorafenib (Bayer, Leverkusen, Germany) submitted a simple state-transition model with three health states: PFS, PD and death. This model compared sorafenib with BSC for people who had failed treatment with immunotherapy or were not suitable to receive immunotherapy. Data from TARGET,[91] a Phase III multicentre placebo-controlled RCT assessing the effectiveness of sorafenib, were used to model both treatment arms in the economic model. Owing to the short follow-up and consequent immaturity of the data, survival had to be extrapolated beyond the trial; this was done using an exponential function. The company presented analyses for the overall population and for the two subgroups: patients for whom immunotherapy failed and patients who were unsuitable for immunotherapy but experienced failure on a non-immunotherapy-based first-line treatment. The company also presented an exploratory analysis comparing sorafenib with sunitinib as second-line treatments; however, the subgroup data and indirect comparison were marked as academic in confidence and so only the data for the overall population were presented.

Utilities were elicited from an unpublished survey of physicians, which resulted in values of 0.737 and 0.548 for the PFS and PD health states, respectively. The estimated cost of sorafenib was £2504.60 for a 112-tablet pack of 200-mg tablets.

The results showed an ICER comparing sorafenib with BSC of £90,630 per QALY gained for the overall population in the base case. All OWSAs produced ICERs that were at least £60,000 per QALY gained, with the most sensitive parameters being the HSUVs for PFS and PD, and the resources associated with the number of inpatient days required when receiving sorafenib and BSC.

At the request of the AG, the company submitted a revised cost-effectiveness analysis for the whole trial population and for 83% of trial participants in whom immunotherapy had failed. The revised analysis also incorporated a patient access scheme (PAS), which is a mechanism for a company to provide its drug to

the NHS in a more affordable way when there is a large degree of uncertainty in the drug being cost-effective. These PASs can be a simple discount on the list price or a more complex scheme, for example, only a fixed number of doses being funded by the NHS. These schemes are agreed between the company and the Department of Health and are often commercial in confidence (CiC). The PAS included in the revised analysis was a complex PAS, whereby the first pack of sorafenib was free to the NHS. A new price for a pack of 112 tablets (200 mg) was used, which amounted to £2980.47. In the new analysis, PFS and OS curves were modelled using a Weibull distribution instead of exponential functions. The company also made changes to the assumptions around cost and utilities. The resulting ICER for the overall population was £72,546 per QALY gained. The ICER for the subgroup in whom immunotherapy had failed was £62,256 per QALY gained. No SAs were presented for the revised analysis. The company also performed an analysis for the subgroup of 17% of participants in whom other first-line treatments had failed, but this analysis was marked confidential.

Company's submission: sunitinib (Pfizer Inc., NY, USA)

The holder of the marketing authorisation for sunitinib (Pfizer Inc., NY, USA) submitted a simple state-transition model with three health states: PFS, PD and death. The model was used to assess the cost-effectiveness of sunitinib compared with BSC (defined as monitoring of progression, symptom control and palliative care without active treatment) as second-line treatments in RCC. Data for the effectiveness of sunitinib were taken from a single-arm Phase II trial with 63 participants who experienced progression on cytokine therapy, while BSC data were taken from a pooled analysis of a systematic review by Motzer *et al.*[100] and an analysis of data from the SEER programme linked to Medicare data.[92,97,100] Disease progression, survival and treatment effect were modelled using survival analysis. Weibull survival curves were used to extrapolate independent data from different sources.

Health state utility values were based on EQ-5D data collected in the single-arm Phase II trial evaluating the effectiveness of sunitinib.[92] Different values were assigned to the PFS health state for each of the two treatment arms. PFS on sunitinib had a utility of 0.803; PFS on BSC had a utility of 0.758. PD on both sunitinib and BSC had a utility of 0.683.

The cost-effectiveness estimates incorporated a PAS in which the first pack of sunitinib was free to the NHS. The resulting ICER for sunitinib compared with BSC was £37,519 per QALY gained in the base-case analysis. OWSAs showed that the ICER was most sensitive to time spent in progression and the source of the effectiveness data used for BSC. The ICERs ranged from £27,935 to £206,962 per QALY gained when these parameters were tested.

Assessment group model

The AG developed a Markov model evaluating the comparative cost-effectiveness of second-line sorafenib and BSC. The model used three health states: PFS, PD and death. Baseline disease progression was modelled by fitting Weibull curves to the empirical PFS and OS curves from the BSC arm of TARGET.[91] Disease progression in the sorafenib arm was estimated by applying HRs from TARGET. The cost-effectiveness of sunitinib compared with BSC was not assessed as the data only came from a single-arm trial.

Health state utility values used in the model were based on trial data from the company (Pfizer) submission and UK EQ-5D tariffs. Treatment-specific utilities were not applied as it was assumed that patients had similar HSUVs at baseline. Utilities for second-line treatments were 0.76 in the PFS state and 0.68 in the PD state.

Drug list prices were taken from the *British National Formulary* (BNF)[101] and the PAS for sorafenib was applied.[102] All other costs were updated to 2007–8 values using the same updated sources. Additional resource use associated with outpatient monitoring, scans and tests were used in the model for people in the PFS state on drug treatment. In the PFS state, the medical management cost per cycle was £81 for BSC and £223 for sorafenib. In the PD state, the cost for each cycle was £435 for both treatment arms.

A number of one-way and multiway SAs were performed to test the sensitivity of the results to variation in model inputs. These included the variation of assumptions made on clinical effectiveness, drug acquisition and administration costs, BSC and management costs, and HSUVs. The AG highlighted in particular a paucity of data around HSUVs and BSC costs.

The results of the AG's model produced an ICER of £102,498 per QALY gained using the original price of sorafenib. SAs showed that treatment effectiveness and the cost of sorafenib were key drivers of the ICER. The ICER was particularly sensitive to variations in the HR for OS, varying from £55,585 (HR 0.54) to £368,830 (HR 0.94) per QALY gained.

As already mentioned, Bayer provided a revised analysis of the whole trial population and of the 83% of participants for whom first-line immunotherapy had failed. No revised analysis for the 17% of patients for whom other first-line treatments had failed was presented as it was confidential. The revised analysis also included details of a complex PAS, whereby the first pack of sorafenib was provided free to the NHS. An increased drug price was also given in the revised analysis.

The NICE Decision Support Unit (DSU) was asked to appraise the approach used by the company in their revised analysis and provide cost-effectiveness estimates using the AG's model, incorporating the PAS and the new increased price. The DSU agreed with the company that the PH assumption was not valid; this resulted in a large reduction in the ICERs. The DSU also noted that the revised analysis presented a lower ICER in people who failed immunotherapy than in the overall population. This was reported as substantially different from the original (confidential) analyses.

The DSU modelled the PFS and OS curves for sorafenib and BSC using independent Weibull curves. The revised ICER for the whole population was £74,915 per QALY gained. The revised ICER for the subgroup of participants for whom immunotherapy had failed was £65,929 per QALY gained.

Single technology appraisal number 219: everolimus for the second-line treatment of advanced renal cell carcinoma

The holder of the marketing authorisation for everolimus (Novartis Pharmaceuticals, Basel, Switzerland) presented a cost–utility analysis assessing the cost-effectiveness of everolimus plus BSC compared with BSC alone in patients with advanced RCC. The company developed a Markov model consisting of four health states: stable disease without AEs, stable disease with AEs, PD and death. The model had 8-weekly cycles and a time horizon of 144 weeks, which was justified as being the maximum life expectancy of the population in an analysis of the RECORD-1 trial, used as the main source of clinical evidence for the economic evaluation.[93] The target population of the model was defined as adults with advanced RCC whose cancer had progressed within 6 months of receiving sunitinib and/or sorafenib, and had demographic characteristics reflecting those of the RECORD-1 trial. Patients in the trial were allowed previous therapy with a cytokine or bevacizumab.

Transition probabilities from the initial state of stable disease without AEs were calculated using the incidences of grade 3 and 4 AEs, treatment withdrawal, disease progression and death obtained from the RECORD-1 trial. As the trial provided data only up to the seventh cycle in the model (i.e. 56 weeks), the company assumed that the event rates remained constant after this cycle until the end of the time horizon. Transition probabilities to stable disease with AEs and PD were calculated directly from event frequencies observed in each arm of the trial. For transitions to death, this method was used only for the everolimus plus BSC arm of the model, while a HR was estimated and applied to calculate the transition probabilities for the BSC arm. Owing to the presence of crossover between the everolimus and placebo arm in the RECORD-1 trial, the company adjusted the HR for OS using the inverse probability of censoring weighted (IPCW) method. The rationale for applying a HR to estimate BSC transition probabilities for death, rather than using the trial data directly in the same way as the everolimus arm, was to keep a constantly higher relative risk of mortality for any given cycle in the BSC model arm compared with the everolimus arm.

The HSUVs were taken from the sunitinib company submission in a previous MTA in the same disease area (TA178). The sunitinib model included utilities for a PFS state and a PD state. In the everolimus model, the PFS utility was applied to the 'stable disease without AEs' health state and the PD utility score was used for the 'progressed disease' state. For the stable disease with AEs, a disutility of 0.05 was assumed, regardless of which AE(s) occurred. This disutility was applied for the first cycle only when patients entered the stable disease with AEs, as AEs were assumed to be resolved within one cycle.

Treatment costs were included in the model as per the proposed PAS, whereby the first treatment pack (30 tablets, 10 mg each) was free to the NHS while subsequent treatment packs were discounted by 5% to £2822. Costs associated with BSC, monitoring and AEs were taken from TA178.[29] Subsequent treatment costs were also included after everolimus treatment had ended, including sunitinib, sorafenib and bevacizumab. Results of the model were later updated with a revised PAS which was designated as CiC. This report presents the initial and revised ICERs.

In the original analysis, the base-case ICER was estimated to be £51,613 per QALY gained (see *Appendix 12*). OWSAs showed that the ICER was most sensitive to the OS estimate in the BSC arm. The PSA estimated that the probability that everolimus plus BSC was cost-effective at a threshold of £50,000 per QALY gained was 40%. An additional analysis was performed by the company using the ITT OS HR (i.e. not adjusted for crossover). In this SA, the HR changed from 0.55 in the base-case analysis to 0.87, resulting in an ICER equal to £91,256 per QALY gained.

The ERG commented that the clinical effectiveness evidence was of good quality, albeit from just one RCT being found through the systematic review. The ERG found that the OS estimate was the main factor affecting cost-effectiveness results, and it agreed that it was important to adjust the data for confounding due to crossover in the trial. The ERG identified and corrected two technical errors in the estimation and application of the HR for mortality in the model and in how discounting was applied. Correcting the two errors increased the company's base-case ICER of £51,613 to £65,231 per QALY gained. The ERG noted that other methods used to adjust for treatment crossover such as the RPSFTMs could have been investigated to assess the impact of adjusting for crossover in the economic results.

The company produced analyses using the RPSFTM in response to the ERG's comments, which resulted in an ICER of £53,128 per QALY gained (with all other base-case assumptions unchanged). However, the ERG felt that the mortality risk in the BSC arm had been overestimated due to an extrapolation based on a single data point. The ERG felt that more data should have been used in the RPSFTM analysis.

Following the revision to the PAS and the ERG's comments, the company provided an updated cost-effectiveness analysis, incorporating the ERG's assumptions in the RPSFTM analysis, more recent trial data (November 2008 cut-off point) and a longer time horizon. The ICER from this analysis was £49,272 per QALY gained. The ERG considered the changes reasonable but reiterated their concerns about the wide CIs for the OS HR, particularly as OWSAs showed that the ICER was associated with substantial uncertainty.

A PSA was performed, which resulted in a mean ICER of £50,047 per QALY gained. The company believed that the wide CI for the OS HR was not clinically plausible, and performed an additional analysis with an adjusted range for the PSA. This resulted in an ICER of £47,811 per QALY gained.

The ERG noted that it was unclear whether or not all sources of uncertainty had been included in the company's PSA. When the ERG re-ran the PSA, keeping the adjusted CI for the OS HR, it estimated the mean ICER to be £49,479 per QALY gained. With the original CI, the ICER increased to £51,661 per QALY gained in the ERG analysis. The results of the final ERG base-case analysis are shown in *Appendix 12*.

Update of TA219[88] (TA432[26])

A rapid review of TA219 was carried out to reassess everolimus for the second-line treatment of advanced RCC in the UK, as it was not available outside of the CDF. The updated guidance was published in February 2017. The company submitted a model in line with the appraisal committee's recommendations set out in TA219 and incorporated a confidential PAS, which was a simple discount on list price.

The model included updated costs as the estimates used in TA219 were deemed to be out of date. The company also analysed effectiveness data using a RPSFTM instead of the IPCW used in the original analysis in TA219. The ERG did not review the RFSTM, but reported that the company's results for the new base-case analysis were very similar to the ERG's proposed ICER in TA219.

Furthermore, the ERG had issues with some changes in parameters compared with TA219, with no justification provided. For example, the HSUV for PD was changed to 0.36 and the dose intensity for everolimus was assumed to be 88%, with both values not being consistent with TA219. The ERG identified errors in the PSA; one was a programming error which when corrected gave similar results to the company's PSA. However, there were additional issues in the choice of parameters that were varied, which limited the usefulness of the PSA results, and the ERG considered that it potentially underestimated the uncertainty surrounding the results.

The company also submitted an exploratory analysis, which included axitinib as a comparator, as axitinib is the current standard of care and not BSC. However, only key results were presented for review without a model. Effectiveness data from the AXIS and the RECORD-1 trials were used to inform the exploratory analysis.[43,64] The company carried out a matched adjusted indirect comparison with alignment of the patient populations in the two trials, using an approach similar to propensity score weighting, in order to estimate PFS for axitinib and everolimus that could be compared as if they had been in the same study. Individual patient data (IPD) from RECORD-1 were used to perform the matched adjusted indirect comparison with the AXIS cohort, using summary outcome measures owing to the lack of available IPD. The company concluded that everolimus had a slightly better PFS than axitinib, which the ERG disagreed with based on other published analyses by the company that showed that everolimus and axitinib had very similar PFS.[103] The company assumed OS for everolimus and axitinib to be the same and based it on OS observed in RECORD-1. The ERG considered this to be in line with clinical expert opinion. The ERG highlighted the same issues in terms of the choice of parameters applied for costs and utilities in the updated analysis were also applied in the exploratory analysis.

The ICERs in both the updated and exploratory analysis are confidential.

Single technology appraisal number 333: axitinib for treating renal cell carcinoma after failure of prior systemic treatment

The holder of the marketing authorisation for axitinib (Pfizer) presented a cost–utility analysis that evaluated axitinib compared with BSC in people with advanced RCC after failure of treatment with sunitinib or a cytokine. The analysis used data from three RCTs of second-line treatments for amRCC: AXIS, RECORD-1 and TARGET.[43,64,104] The AXIS trial compared axitinib with sorafenib, TARGET compared sorafenib with placebo, and RECORD-1 included a placebo plus BSC arm following first-line sunitinib treatment. These studies were identified by systematic review and were used to form an indirect comparison of axitinib with BSC due to the lack of head-to-head RCT evidence. Subgroup analyses by prior treatment were carried out.

For the prior-cytokine therapy group, a MTC was performed using a Markov chain Monte Carlo simulation in WinBUGS to estimate the relative treatment efficacy of axitinib and BSC, using data from the placebo arm of TARGET, under the assumption of equivalent effectiveness to BSC following cytokine therapy. HRs for both PFS and OS from AXIS and TARGET were used in a fixed-effects model assuming PHs. Point estimates and 95% CrI were estimated and the results showed a PFS of 11 months for the axitinib group compared with 3.5 months in the BSC group (median HR 0.25, 95% CrI 0.17 to 0.38). The axitinib group

showed a median OS of 33.5 months compared with the 23.5 months in the BSC group (median HR 0.63, 95% CrI 0.41 to 0.99). The HR for OS was based on data censored at crossover.

The company performed a simulated treatment comparison (STC) for the first-line sunitinib arm, as there was no direct or indirect trial evidence This compared PFS with OS for axitinib versus everolimus and BSC using data from the AXIS[43] and RECORD-1[93] trials. Owing to several differences between the patients in the trials, two approaches were tested to simulate the treatment relative effectiveness. The first was by comparing the ITT placebo group in RECORD-1, which included patients who have received sunitinib and/or sorafenib, with the sunitinib-refractory patients in AXIS. This implicitly assumes that the patients who received prior sunitinib and prior sorafenib have similar characteristics. The second approach compared the axitinib post-sunitinib group with the everolimus post-sunitinib group and then applied the RPSFTM-adjusted HR for everolimus to BSC to create a modelled prior-sunitinib group.

The simulated relative treatment effectiveness was performed by analysing IPD from the axitinib arm of AXIS to derive parametric survival functions incorporating baseline predictors of the two main trial outcomes (i.e. PFS and OS). Five different distributions were examined, but only the two best-fitting models were used for both PFS and OS, which were the log-normal and Weibull distributions. The results of the analysis showed a benefit for axitinib relative to BSC and everolimus in both PFS and OS, when log-normal and Weibull distributions were used. However, the data relating to the estimated increase in PFS and OS in the ITT population and prior-sunitinib subgroup were CiC.

The PFS HR of 0.34 for the prior-sunitinib group and the adjusted OS HR of 0.53 for the ITT group of RECORD-1 were applied to the everolimus STC curves to generate AXIS-like prior-sunitinib PFS and OS curves for BSC. This resulted in an estimated median PFS of 1.7 months for the 'axitinib-like patients' if they had received placebo, compared with a median of 5.8 months if they had received axitinib (HR not reported). The median OS estimated for the same patients was 8.3 months for placebo compared with 15.2 months for axitinib (HR not reported).

The company also provided an additional analysis that used retrospective observational data for patients who received first-line sunitinib followed by sorafenib or BSC from a Swedish database, Renal Comparison (RENCOMP), to estimate the OS HR for people who received sorafenib or BSC following first-line sunitinib treatment.[105] A multivariate Cox PH regression analysis was performed using variables with significance at the 5% level to account for any bias resulting from confounding. The results of this analysis showed a median OS HR of 0.62 (95% CI 0.41 to 0.94; $p = 0.023$). The results from RENCOMP were used in an indirect comparison with the results from the first-line sunitinib group in the AXIS trial, which compared axitinib with sorafenib (median PFS HR 0.74, 95% CI 0.57 to 0.96; median OS HR 0.997, 95% CI 0.78 to 1.27), to estimate indirect HRs for axitinib compared with BSC in the prior-sunitinib group. The resulting HR was 0.62 (95% CI 0.38 to 0.997).

The company developed a Markov model for the cost–utility analysis consisting of three health states: PFS, PD and death. All patients were assumed to enter the model in the PFS state and could transition to any of the other states or remain in the current state after each 4-week cycle. The time horizon of the model was 10 years.

Parametric survival functions were used to calculate the proportion of patients in each health state at each point in time. For the first-line cytokine group who received axitinib at second line, Weibull models were used to extrapolate PFS and OS as this was considered to have the best fit among the alternatives tested. Log-logistic and Gompertz distributions were also tested in SAs for OS, and the log-normal and Gompertz models were used alternatively in SAs for PFS. For the purpose of estimating PFS and OS in the BSC arm, HRs estimated from the indirect comparison analysis were applied to the parametric survival functions used for the axitinib arm.

Utilities were taken from the AXIS trial using the EQ-5D questionnaire analysis undertaken in the trial. The analysis was based on the full population, as there was no statistically significant difference in utilities between patients receiving different types of first-line therapies. The utility for the PFS health state was 0.69 and was 0.61 for the PD state. The PFS utility was derived from the average EQ-5D index value at each time point, weighted by the number of people still on treatment, whereas the PD score was taken as the weighted average of the mean utility at the end of treatment. These utilities were also assumed to incorporate implicitly the AE profile of the treatments in the trial. A systematic review of the literature did not find any sources of utilities for patients with amRCC receiving BSC after sunitinib treatment had failed. Therefore, the utilities from AXIS were also assumed to apply to BSC. Utilities from previous NICE TAs were explored as part of the SAs.

Costs of treatment were based on the proposed PAS, which was CiC. Costs were adjusted for dosing intensities and discontinuation probabilities were applied at each cycle during treatment (0.8% and 1.26% for prior-cytokine and prior-sunitinib, respectively) while AE rates were assumed to be the same between subgroups. AE-related costs were included in the PFS health state only and assumed to be equal across the two prior treatment subgroups. Only grade 3 and 4 AEs that occurred in > 5% of the population were included. For axitinib, hypertension had a cost of £424.00 per episode and diarrhoea had a cost of £544.00 per episode. In the BSC arm, anaemia was included at a cost of £2068.47 per episode as a SA.

No administration costs were incurred and no drug costs for the BSC arm were assumed to apply. Costs of routine monitoring were based on TA178 and were validated with expert clinical opinion to ensure consistency with current clinical practice.[29] These costs were assumed to apply to both arms of the model equally. The total cost per cycle for the PFS state was £109.69 based on one GP visit, one tumour scan per three cycles and one blood test per cycle. For the PD state, the cost per cycle was £319, based on one GP visit, three visits by a specialist community nurse every two cycles and 28 vials of pain medication per cycle. A SA explored changing the GP visit to an oncologist visit, which changed the costs for the respective health states to £176.69 and £386.00.

The updated economic analysis increased the time horizon to 15 years after the ERG's concern that 10 years might not have reflected a lifetime horizon in the economic analysis. First-line cytokine and first-line sunitinib subgroup specific utilities and relative dose intensities (RDIs) were applied rather than estimates for the ITT population used in the original model. The percentage of people with hypertension was reduced from 2% to 0% as the percentage of hypertension in TARGET was < 1%.[104] The PSA was updated to include SEs rather than SDs for the PFS and PD HSUVs, RDI and the cost of death. An error in the STC was also corrected, which had only a marginal effect on the results.

The results of the first-line cytokine subgroup analysis showed an ICER of £55,284 per QALY gained. A range of OWSAs were performed using the ranges of the 95% CIs, and these showed that the results were most sensitive to the HR of OS and the PD utilities. ICERs ranged from £40,000 to £100,000 per QALY gained for changes in utilities and up to £350,000 per QALY for changes in the OS HR. The results of changing the method of extrapolation for OS varied from £21,959 per QALY gained for the log-logistic method to £72,537 per QALY gained for the Gompertz method. Other scenario analyses resulted in ICERs similar to the base-case analysis.

For the first-line sunitinib subgroup the base-case ICER was £33,538 per QALY gained. OWSAs showed that the results were most sensitive to changes in the parameters for the survival functions for PFS and OS, with ICERs ranging from around £25,000 to £48,000 per QALY gained. The ICER was also sensitive to the PD HSUVs for both arms and for PFS in the axitinib arm. The resulting ICERs ranged from around £29,000 to £40,000 per QALY gained. Other OWSAs had very little impact on the results. The ICER was most sensitive to different survival distributions using the method of comparison based on RENCOMP data, with use of the Weibull and Gompertz distributions resulting in ICERs of £47,515 and £39,479 per QALY gained, respectively. Reducing the RDI to 80% resulted in an ICER of £27,324 per QALY gained.

The company also applied an assumption of no QALY or survival gain post progression, which resulted in an ICER of £52,850 per QALY gained.

Before the end of the STA process, the scope of the appraisal was updated by NICE after an appeal hearing, with sunitinib and pazopanib added as comparators for the prior-cytokine subgroup. The company conducted an additional systematic review to search for relevant evidence to inform the new comparisons. As no head-to-head trials were found, indirect comparisons and naive indirect comparisons (assuming trial groups are homogenous) were used to estimate relative treatment effects. In addition to AXIS, two trials were identified. The first trial compared pazopanib with placebo[106] while the second was a single-arm trial assessing the efficacy safety of sunitinib.[107] A naive comparison between axitinib and sunitinib showed that axitinib increased PFS by 3.3 months and OS by 5.5 months. A naive comparison between axitinib and pazopanib showed that axitinib increased PFS by 4.7 months and OS by 6.7 months. An indirect comparison showed that PFS was longer for axitinib compared with pazopanib (median HR 0.47, 95% CrI 0.26 to 0.85). The company did not conduct an indirect comparison for OS owing to the crossover between placebo and treatment groups in the axitinib and pazopanib trials.

The company did not provide a full incremental analysis with the new comparators but instead provided a naive economic comparison based on the base-case analysis with the PAS (with an ICER of £55,284 per QALY gained). The naive treatment comparison showed an increase in median OS for BSC (24 months compared with 23.9 months and 22.7 months for sunitinib and pazopanib, respectively), so the company used the lower 95% CI of 17.6 months for a more realistic estimate. The recalculated ICER for axitinib versus BSC was £36,493 per QALY gained and an upper limit of £55,000 per QALY gained was estimated for the ICERs comparing axitinib with either sunitinib or pazopanib.

The company also provided an additional analysis that used data from a retrospective cohort study,[108] showing correlation between tumour shrinkage and OS, to weight the median OS estimates in the BSC arm of RECORD-1 by multiplying the median overall survival by the proportion of patients. The resulting weighted median OS for BSC was 8.2 months (95% CI confidential information has been removed) which the company believed was consistent with the 8.3 months OS result from the STC. The company concluded that the results were consistent with the results of the base-case analysis using the STC, which resulted in an ICER of £33,538 per QALY gained.

Evidence review group's comments

The ERG's critique stated that there were few limitations with the literature search conducted by the company and that the studies found were good-quality clinical trials with robust methodologies, except for the adjustment of crossover in TARGET (censoring at crossover). The ERG considered the RPSFTM approach used in RECORD-1 to be a more appropriate adjustment method for crossover. The ERG also noted that the patient characteristics were not reported separately for the first-line cytokine subgroups in either AXIS or TARGET, so the ITT populations were used for the indirect comparisons. The ERG considered that the populations were reasonably similar in the two trials, but considered that the potential bias associated with the HR for OS in TARGET (due to the crossover adjustment method used) may limit the robustness of the indirect comparison in the prior-cytokine group.

The ERG was uncertain about the validity and reliability of the STC as it used data from two treatment arms from two different RCTs, thus breaking randomisation and potentially introducing bias. The results could not be verified by the ERG, as IPD were not provided. The ERG also noted the use of observational data from RENCOMP was a potential source of bias due to the lack of randomisation in treatment allocation and the reasons for discontinuation being unknown.

The ERG was satisfied with the modelling methods used, which were consistent with other published economic studies of advanced RCC. The ERG agreed with the company's assumption that the HSUVs were the same for people receiving axitinib and BSC. However, the ERG considered that the HSUVs should not remain constant in the PD state but should decline as patients neared the end of life. The ERG expected

this change to increase the ICER slightly. The ERG also noted that the AXIS clinical trial report stated that the US valuation was used for utility estimation, and noted that the US valuation usually yields consistently higher values than the UK one. Finally, the ERG pointed out that surrogacy estimates between PFS and OS, which the company used for validation of the results, had changed significantly in the subsequent publication.

Following the appeal hearing and the consequent updated scope, the ERG agreed that an indirect comparison was not possible between axitinib and sunitinib. The ERG considered the ITT populations of the trials used for the indirect comparison of axitinib and pazopanib to be reasonably comparable, but the patient characteristics in the first-line cytokine subgroup were not presented separately. The ERG conducted additional analyses for an indirect comparison of OS between axitinib and pazopanib, using IPCW and RPSFTMs to adjust for crossover in the placebo arm after progression. In the ITT population, the HR was 0.77 (95% CrI 0.44 to 1.38), in favour of axitinib. The HR from the IPCW analysis was 1.20 (95% CrI 0.55 to 2.61) and in the RPSFTM analysis the HR was 1.21 (95% CrI 0.30 to 4.82), with both in favour of pazopanib. None of the results showed statistical significance and the ERG stated that none of the indirect comparison results was likely to be reliable because the common treatment effect assumption was unlikely to apply for the RPSFTM analysis, the no unmeasured confounder assumption was unlikely to apply to the IPCW analysis and, therefore, the results from these analyses were likely to be biased. The ERG stated that the PFS was most likely to be reliable because it was unaffected by crossover and there was no clear evidence of difference between treatments.

The ERG commented on the limitations of the naive economic analysis, stating that it was only an indicative analysis and it did not provide an estimate of the cost-effectiveness of axitinib compared with sunitinib or pazopanib. The ERG also found an error in the company's submission; after the corrections, the ICER was £36,493 per QALY gained and not £33,000.

The ERG stated that the real OS for BSC was not known owing to a lack of trial data, so the additional analysis carried by the company using the surrogate end point of tumour shrinkage to estimate OS in the BSC arm of the model could not be validated. The ERG noted that the study evaluating the correlation between tumour shrinkage and OS did not include patients receiving BSC and, additionally, that it was not clear how tumour shrinkage was assessed in the different trials in the analysis, whether or not there were any other factors correlated with either tumour shrinkage or OS or both, and whether or not crossover between treatments was permitted after treatment failure.

Single technology appraisal number 417: nivolumab for previously treated advanced renal cell carcinoma

The holder of the marketing authorisation for nivolumab (Bristol-Myers Squibb, NY, USA) submitted a cost–utility analysis evaluating the cost-effectiveness of nivolumab in comparison with everolimus, axitinib and BSC. A six-state transition model was developed with health states defined as: PFS on treatment, PFS off treatment, post-progression survival on treatment, post-progression survival off treatment, terminal care (TC) and death. All patients started in the PFS on treatment state and could only transition to death via the tunnel state TC, which they were assumed to stay in for 8 weeks before death. The time horizon was set at 30 years and weekly cycles were used.

Model parameters were informed by data from the CheckMate 025 trial,[54] a Phase III multicentre open-label RCT comparing nivolumab with everolimus in patients with histologically confirmed advanced/metastatic RCC who have received one or two previous antiangiogenic agents. A MTC including nine studies was performed to estimate the treatment effects on OS and PFS between nivolumab and axitinib, and nivolumab and BSC.[54,66,91,93,109–113]

To extrapolate survival for nivolumab and everolimus, a generalised gamma model was used, which was considered to give plausible results based on expert clinical opinion. The relative effectiveness of everolimus and BSC was incorporated by applying the crossover-adjusted HR from the network meta-analysis (NMA) to the OS curve of the everolimus arm data from CheckMate 025. It was assumed that BSC would be as effective as placebo.

The company considered that standard parametric models did not fit the PFS data sufficiently well so more flexible spline-based models were explored. Models from Royston and Parmar[114] were applied to the PFS data from CheckMate 025. The PFS data for axitinib and BSC were estimated by applying the HR from the NMA analysis to the everolimus survival curve, assuming that BSC was as effective as placebo.

The company used the same spline-based survival analysis approach to model time-to-discontinuation (TTD) data for nivolumab and everolimus as for PFS. In the absence of TTD data for axitinib, the company assumed that treatment was continued until disease progression.

Pharmacological resource use was based on the treatment indications and adjusted for dosage, with 92% and 94% of the planned dosage being costed for nivolumab and everolimus, respectively, based on the CheckMate 025 trial. For axitinib, the actual usage was 102% of that planned based on the AXIS trial, as reported in TA333.[25] Resource use in the PFS and post-progression survival (PPS) states was assumed to be the same as that in TA333. The cost of TC was considered in the model and based on a paper by Addicott and Dewar on improving choice at the end of life.[115]

The company only included serious grade III/IV treatment-related AEs experienced by $\geq 1\%$ of patients in either arm of the CheckMate 025 trial. These were pneumonitis, anaemia, diarrhoea and pneumonia. Rates and duration of these AEs were based on the trial observations for both nivolumab and everolimus, and management costs were applied weekly in the model. Owing to a lack of data, the cost of managing AEs for patients treated with axitinib was assumed to be equal to that of everolimus treatment. The weekly cost of AEs was £0.35 for nivolumab and £1.31 for axitinib and everolimus.

The HSUVs were applied to each health state and were derived from EQ-5D data obtained from the CheckMate 025 trial. A mixed-effects model was used to analyse the data, with fixed effects for progression status and treatment allocation; a variable effect for the interaction between treatment arm and progression status; and a random effect for the subject. For patients receiving axitinib, HSUVs were derived from EQ-5D data from the AXIS trial, reported in TA333. It was assumed that patients receiving BSC would experience the same quality of life as patients receiving axitinib. The HSUVs before progression were 0.80, 0.76, 0.69 and 0.69 for nivolumab, everolimus, axitinib and BSC, respectively. For PPS, the HSUVs were 0.73, 0.70, 0.61 and 0.61 for these treatments, respectively.

The results of the company's model showed an estimated survival benefit of 16, 11 and 24 months (undiscounted) for nivolumab compared with axitinib, everolimus and BSC, respectively. Nivolumab increased mean (discounted) QALYs by 1.07, 0.63 and 1.43 compared with axitinib, everolimus and BSC, respectively. The company estimated pairwise ICERs of £42,417, £83,829 and £56,427 per QALY gained for nivolumab compared with axitinib, everolimus and BSC, respectively.

Evidence review group's comments

The ERG considered the results of the NMA used to provide a comparison between nivolumab and axitinib unreliable. This was due to the included trials having a wide range of different prior treatments, inconsistent use of adjustments for crossover for estimating OS, and the use of immature OS data from one important link in the network. The ERG considered this weakness to be the most uncertain aspect of the economic analysis. The company's base-case analysis was not considered plausible by the ERG's clinical experts and or by the oncologists interviewed by the company. In the opinion of the ERG, the company did not convey the substantial uncertainty associated with the relative treatment effectiveness estimates between everolimus and axitinib, none of which showed a statistically significant difference between the two treatments in the company's NMA.

The uncertainty associated with the relative treatment effects was increased by methodological errors in applying HRs derived from the NMA to non-PH survival models. This produced relative effectiveness estimates considered unreliable by the ERG, in particular between nivolumab and axitinib, and nivolumab and BSC.

Although a comprehensive set of models was tested in the survival analysis, the ERG considered that there was a lack of testing of the assumptions of the different models. Given the uncertainty in treatment effects, and the reliance of long-term estimates on the parametric models, the robustness of the OS, PFS and TTD projections over the 30-year time horizon should have been further tested according to the ERG.

The ERG was not satisfied with the company's justification of the difference in HSUVs observed in the AXIS trial in comparison with the much higher values obtained from CheckMate 025. Clinical experts agreed that quality of life for patients who progress on axitinib and on everolimus would be comparable. They also agreed that it was unreasonable that patients who did not progress on axitinib treatment would have a lower quality of life than those who did progress on everolimus treatment.

The ERG reported that the proportion of planned drug doses received by patients was not satisfactorily described or justified by the company and that costs included subsequent therapies beyond second line, which are not currently reimbursed in England. These costs should not have been included as the perspective of the economic analysis is that of the NHS in England.

The ERG also noticed several modelling errors but these had very little impact on the company's base-case results. In particular, flaws in the integration of OS, PFS and TTD curves, resulting in negative proportions of patients in the health states or total proportions of patients in health states exceeding 100%. Amendments to the model resulted in increases in the ICERs for the pairwise comparison of nivolumab with axitinib, everolimus, and BSC of £692, £2307, and £331 per QALY gained, respectively.

As well as correcting the errors in the model, the ERG carried out scenario analyses to test the uncertainty around the company's assumptions not sufficiently explored. The first of these was to assess the impact of the relative treatment effectiveness between axitinib and everolimus. This was done by assuming an equal effectiveness for PFS and OS, as clinical experts consulted by the ERG considered axitinib to be at least as effective as everolimus. The ERG also tested the parametric model used by comparing the results with those produced by a log-logistic model to extrapolate OS data from CheckMate 025. This model had the best relative fit to the data among the parametric models tested. As the company did not justify the use of a complex spline-based model for TTD, the ERG also tested a simpler accelerated failure time model instead. This was done using the log-normal model initially but the results using a generalised gamma model are also presented in the ERG's base-case analysis.

The HSUVs were tested by using the same values for axitinib and BSC as for everolimus, as the values from the different trials were inconsistent. This was considered reasonable by clinical experts that were consulted by the ERG.

The proportion of planned drugs received were tested by including the delayed doses of nivolumab in the total doses received. A second scenario was also tested that assumed that patients would receive all planned doses of nivolumab and everolimus. Subsequent therapy costs were also explored by removing these completely, as there are no approved nor reimbursed third-line treatment options for amRCC in England.

The results of the ERG's base-case analysis showed ICERs of £74,132, £91,989 and £61,317 per QALY gained for nivolumab compared with axitinib, everolimus and BSC, respectively. The ERG also explored an equally plausible scenario by using a generalised gamma for TTD, which resulted in ICERs of £81,696, £96,107, and £64,869 per QALY gained for nivolumab compared with axitinib, everolimus and BSC, respectively.

The ERG highlighted substantial uncertainty on the relative treatment effectiveness between everolimus and axitinib and thus between nivolumab and axitinib. The ERG considered that the company did not analyse appropriately the adjustments made to the relative treatment effects because of the presence of treatment switching in the trials included in the MTC. The assumption of equal effectiveness made by the ERG was deemed to be more plausible by clinical experts but it is likely to underestimate the effectiveness of axitinib and hence underestimate the ICER comparing nivolumab with axitinib.

Single technology appraisal TA463: cabozantinib for treating advanced renal cell carcinoma in adults who have received at least one prior vascular endothelial growth factor-targeted therapy

The holder of the marketing authorisation for cabozantinib (Ipsen, Paris, France) submitted a de novo economic model that evaluated cabozantinib in two separate cost–utility analyses. The first, a trial-based analysis comparing cabozantinib with everolimus, using effectiveness data obtained solely from the METEOR trial, a Phase III open-label RCT comparing cabozantinib and everolimus in patients with advanced RCC who progressed after previous VEGF receptors TKI treatment.[57] The second was a pairwise analysis of cabozantinib compared with everolimus, axitinib, nivolumab and BSC based on effectiveness data derived from a MTC of trials identified in a systematic review of the literature.[54,66,93,104]

The two analyses used the same partitioned survival model structure, composed of three health states: PFS, PD and death. Patients entered the model in the progression-free state and could transition to PD or death. From the PD state, patients could transition to the death state at each future cycle. The time horizon was set at 30 years and 4-weekly cycles were used.

A log-logistic distribution was chosen to extrapolate observed OS and PFS for both cabozantinib and everolimus in the METEOR trial in the trial-based analysis. In the MTC-based economic evaluation, parametric survival curves were estimated and extrapolated based on regenerated KM data from CheckMate 025, AXIS, RECORD-1 and TARGET, in addition to KM data from METEOR.[54,57,66,93,104] The NMA estimated the parameters of independently fitted survival curves for each treatment group in each trial in the network and the parameters were adjusted to the everolimus group of the METEOR trial.[57] A single family of distributions was fitted to each group in each trial. The log-normal distribution was deemed to be the best fit for both OS and PFS in the NMA-based analysis and was used in the company's base case for the curves of each comparator.

The TTD data were used to estimate pharmacological costs in both analyses. In the trial-based analysis TTD KM data from METEOR were used and extrapolated using a log-normal distribution. In the NMA-based analysis regenerated KM data from trials were used to estimate the parameters of the best-fitting curve and adjust them to the everolimus group of the METEOR trial. KM data for axitinib were not available and PFS data were used as a proxy for TTD.

Relative dose intensities were applied to estimate the true cost of the treatments compared and were taken from the respective trial data. Treatment administration costs were only applied for nivolumab as it is administered intravenously, with no wastage assumed. Disease management costs were estimated based on the expert clinician opinion of oncologists practising in the UK, and varied according to health state. Costs related to AEs were limited to those events that occurred in $\geq 5\%$ of patients in each group of the relevant trials. Treatment-emergent adverse events (TEAEs) were used in the models for all comparators except nivolumab, as only treatment-related adverse event (TRAE) data were identified for it.

The HSUVs were estimated based on a regression analysis of EuroQol-5 Dimensions, five-level version (EQ-5D-5L) collected in the METEOR trial; utility decrements for AEs were applied. The HSUVs used in the model for patients regardless of treatment arm/analysis were 0.817 prior to progression and 0.777 after progression. A disutility of 0.055 was applied to patients when experiencing AEs regardless of treatment arm.

The results of both analyses are not shared publicly, and remain confidential at the time of writing this report. A range of OWSAs and scenario analyses were performed as well as a PSA to test the impact of uncertainty of all relevant parameters on the model results.

Evidence review group's comments

The ERG was satisfied with the electronic model submitted by the company and considered the model structure to be appropriate and in line with previous models used in RCC. The ERG considered that the

company carried out an appropriate range of scenario analysis to explore the uncertainty surrounding the base-case analysis.

The ERG considered that there was a lot of uncertainty surrounding the extrapolation of OS from the METEOR trial for the trial-based analysis. This uncertainty is because the company did not consider the Weibull distribution to fit the KM data, which would have avoided the extended and potentially unrealistic tail of the resulting extrapolated curve when the log-logistic curve is used. This was also the case for PFS extrapolation, but to a lesser extent.

The ERG considered the key weakness of the NMA-based analysis to be the poor fit of the estimated survival curves for PFS and OS as a result of the company's approach of using a single family of parametric curves for all comparators, with goodness-of-fit being assessed to the model globally. The ERG was concerned that this could cause unrealistic differences in the inherent treatment effect as determined by the resultant independent curves.

The ERG's clinical experts considered that the resource use assumed in the model was not reflective of UK clinical practice, suggesting that GP visits should be substituted by consultant visits in the analysis. Furthermore, the ERG considered that assuming no wastage of nivolumab vials in the base-case analysis is not realistic as wastage is likely to occur. The ERG's clinical experts also believed the utility values used in the model to be an overestimate of patients' quality of life, in their experience. At the time of writing this report, this NICE TA was ongoing.

Independent economic assessment

The AG undertook an independent economic assessment of the cost-effectiveness of treatments for advanced RCC. For that purpose, a de novo economic analysis was developed and is presented in this section of the report. The methods employed, data inputs and results of the economic analysis are also presented throughout this section.

Scope

The scope of the independent economic assessment is described in *Table 13* and reflects the decision problem as outlined in *Chapter 2*.

Model structure

A partitioned survival model approach was used in the cost-effectiveness analysis of RCC treatments. The model uses survival analysis to simulate the progression of RCC in a cohort of patients over time and was written in Microsoft Excel. The structure was informed by a systematic review of available literature and clinical expert opinion.

The model includes the following health states: PFS, PD and death. The model structure is presented in *Figure 7*.

Patients are assumed to start the model in the PFS state and can remain progression free in the next cycle or move to either the PD state or die. Patients entering the PD health state can either remain in the PD health state or die. The proportion of patients in each health state at any point in time is based on parametric survival curves for each clinical outcome as per the partitioned survival approach.

All active treatments being modelled have marketing authorisation to be administered to patients beyond progression. Therefore, the model is flexible enough to reflect this for treatments for which TTD data are available (i.e. everolimus, cabozantinib and nivolumab). Patients in these treatment arms could be either on treatment or off treatment within the PFS and PD health states to accurately capture treatment-related costs. All patients receive subsequent treatments after their second-line treatment.

TABLE 13 Summary of independent economic assessment

Element	Overview	Reference section
Population	People with previously treated amRCC who received previous VEGF-targeted therapy.	*Population*
Interventions	The following interventions were considered in the economic evaluation: • axitinib • BSC • cabozantinib • everolimus • nivolumab	*Interventions*
Outcomes	The outcome measures considered include: • OS • PFS • AEs of treatment • HRQoL	*Treatment effectiveness, Adverse events* and *Health-related quality-of-life data selected for the economic analysis*
Model/economic analysis		*Model structure* and *Time horizon and discounting*
Perspective	NHS and PSS	
Assessment of health benefit	QALYs	
Model type	Partitioned survival	
Time horizon	Lifetime (30 years)	
Cycle length	2 weeks	
Discounting	3.5%	

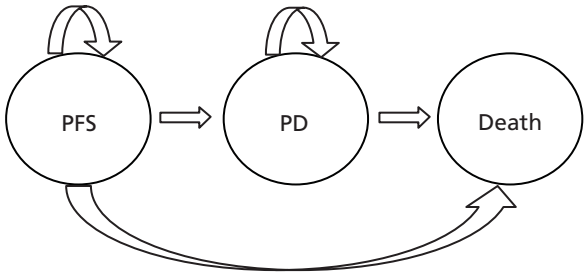


FIGURE 7 Model structure.

The model time horizon is 30 years and the cycle length is 2 weeks. The time horizon and cycle length are considered appropriate for capturing all the relevant health effects, treatment schedules and costs associated with advanced RCC. No half-cycle correction was applied to the model as a result of the short length of the model cycles. The perspective of the analysis is that of the NHS and PSS.

Time horizon and discounting

The time horizon of the economic analysis is 30 years and an annual discount rate of 3.5% is applied to costs and outcomes.

Population

The population of interest for this assessment is people with previously treated amRCC. Based on clinical expert opinion, patients are expected to respond differently to second-line treatment depending on the previous treatment received (i.e. cytokines or VEGF-targeted therapy). However, as cytokines have become

less preferable at first line in the UK, this population is not considered relevant for this analysis, which only focuses on assessing second-line treatments following a VEGF-targeted therapy.

Patients are modelled using the CheckMate 025 trial (and through NMA estimates of relative treatment effectiveness). Therefore, the population in this analysis is reflective of those in the CheckMate 025 trial.

Clinical expert opinion sought by the AG advised that the population in the CheckMate 025 trial appeared to have a better prognosis at baseline than patients in the AXIS trial.[43,54] Furthermore, according to the AG's clinical experts, the population in the AXIS trial is considered more reflective of patients seen in routine clinical practice in the UK. This means that the baseline survival curves and, therefore, all dependent survival curves, may be optimistic in comparison with routine clinical practice.

Interventions

The interventions considered in the cost-effectiveness analysis are axitinib (5 mg, twice daily), everolimus (10 mg, once daily), nivolumab (3 mg/kg, every 2 weeks), cabozantinib (60 mg, daily) and BSC.

Treatment effectiveness

The effectiveness of the treatments compared in the economic model is primarily measured by the rate of disease progression and the rate of death in the population receiving the treatment. Therefore, this section outlines the methods used to calculate the proportion of patients occupying each of the health states, PFS, PD and death, at a given time point for each comparator treatment assessed in the economic model. Owing to the similarity in the methods used to model the rate of treatment discontinuation, this is also covered in this section. As the main focus of this section relates to parametric survival modelling, the general approach taken to fitting survival curves will be discussed first in *Survival modelling*, followed by a description of the specific methods relating to each of the outcomes OS, PFS and TTD.

Survival modelling

The first stage in generating the survival curves was to obtain the data points from published KM survival curves. The images of the KM plots from publications identified in the systematic review were digitised using *g3data* software (version 1.5.1; www.frantz.fi/software/g3data.php), which enabled time points and survival probabilities to be generated and saved as a comma-separated text file. These data, along with the numbers of patients at risk at various time points, was then loaded into *R* version 3.3.2 (The R Foundation for Statistical Computing, Vienna, Austria) to perform the analysis.

The next stage of the process was to generate pseudo IPD using the algorithm derived by Guyot *et al.*[116] This algorithm was implemented using the *R* code in the pre-release version of the *survHE* package, in which this algorithm had previously been applied. The pseudo IPD data were then inputted into the survival functions built into the *flexsurv* package of *R*, to assess and generate fitted survival functions.[117]

Overall survival

The following distributions were fitted to the pseudo IPD data generated from the OS KM plots from the CheckMate 025 trial: exponential, Weibull, Gompertz, log-normal, log-logistic, generalised gamma, generalised *F* and two hazard-based spline functions with 1 knot and 2 knots, respectively. The Akaike information criterion (AIC) and Bayesian information criterion (BIC) were calculated using the standard *stats* package included in *R* and, in combination with a visual plot of the resulting curves, were used to determine the best-fitting distributions for both treatment groups. A single distribution was chosen for all treatment groups following the advice outlined in NICE Technical Support Document 14.[118] The chosen distribution was also restricted by those that allow proportionality of hazards (i.e. exponential, Weibull, Gompertz and the hazard-based spline functions), in order to appropriately apply the HRs derived in the clinical analysis to produce estimated OS curves for cabozantinib and BSC relative to everolimus. Axitinib was assumed to have the same OS curve as everolimus as a reliable analysis could not be conducted to estimate a HR for OS.

The resulting survival curves along with the KM data are presented in *Figures 8* and *9* for nivolumab and everolimus, respectively, and the AIC and BIC statistics are given in *Appendix 15*.

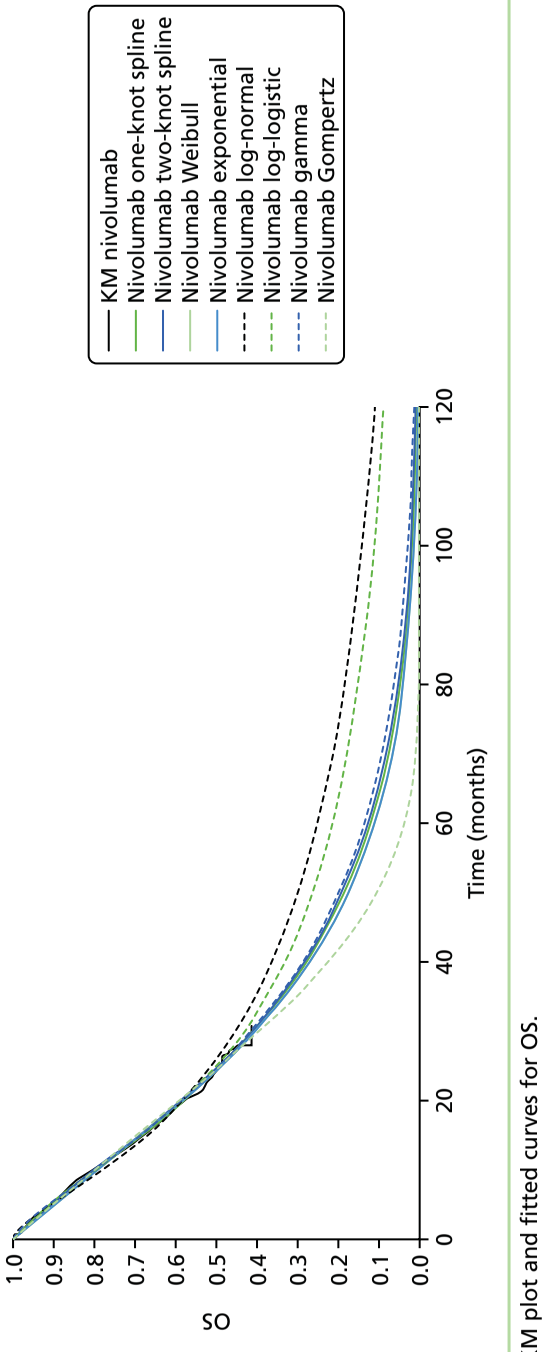


FIGURE 8 Nivolumab KM plot and fitted curves for OS.

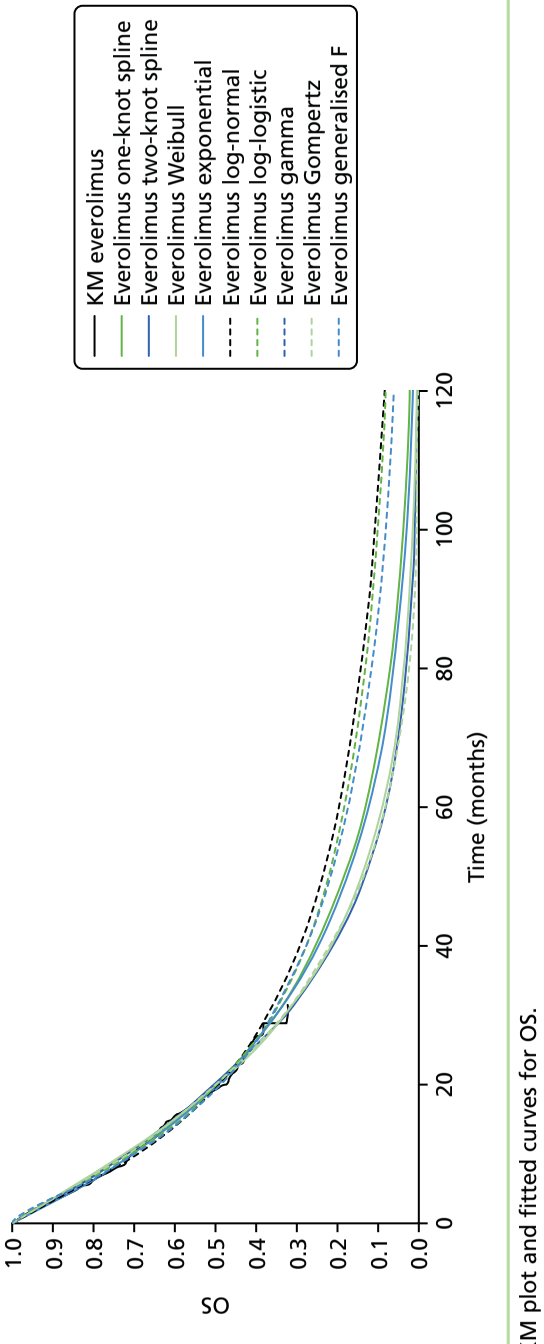


FIGURE 9 Everolimus KM plot and fitted curves for OS.

For nivolumab, the results were unequivocal in favour of the Weibull curve, which had the lowest AIC and BIC statistics, as well as being supported by clinical expert opinion. However, for everolimus, the two-knot spline had the lowest AIC value and the log-normal had the lowest BIC value. Within the subgroup of curves in which hazards may be proportional (exponential, Weibull, Gompertz and splines), the two-knot spline, in contrast to having the lowest AIC, had the highest BIC, so the AG considered that other curves with higher AIC but lower BIC may be equally good fits or potentially better. Given that the Gompertz had higher AIC and BIC values than both the Weibull and the exponential curves, this was considered a less preferable model. The exponential and Weibull models had contrasting statistics in that the AIC for the exponential was higher than that for the Weibull but the BIC value was lower. On visual inspection of these curves, the AG considered the Weibull curve to follow a more plausible projection from the KM data than the exponential curve, and this choice was also in line with the best-fitting distribution for the nivolumab treatment group; thus, the Weibull curve was chosen for nivolumab and everolimus for OS (*Figure 10*). This was supported by expert clinical opinion.

Progression-free survival

The same approach taken for OS was applied to the PFS data from CheckMate 025. The KM plots with fitted curves are given in *Figures 11* and *12* for nivolumab and everolimus, respectively, and the goodness-of-fit statistics are given in *Appendix 15*.

The results of the goodness-of-fit tests clearly indicated that the two-knot spline model had the best fit for both everolimus and nivolumab when assessing the AIC and BIC statistics (*Figure 13*). This curve was validated by clinical experts as having a plausible extrapolation and so this was chosen to model PFS in the base case.

Time to treatment discontinuation

Treatment discontinuation KM plots were identified from the CheckMate 025 trial for nivolumab and everolimus, and from the METEOR trial for cabozantinib and everolimus.[54,57] The CheckMate 025 trial was used as the source of TTD data for the everolimus group to align with the data used for PFS and OS. However, the AG noted the differences between the KM plots for the everolimus groups in the METEOR trial and the CheckMate 025 trial, and considered the appropriateness of using independently fitted curves for the CheckMate 025 trial groups and the cabozantinib group of the METEOR trial, or if adjustment of the cabozantinib curve would be more appropriate. A method of adjustment was used in the submission for the cabozantinib NICE TA (TA463)[28] by performing a NMA to fit independent curves, with a treatment group adjustment applied to the distribution parameters. However, in the AG's opinion, this method may overestimate the rate of discontinuation for the treatment group that is adjusted (in this case cabozantinib) and, therefore, underestimate the treatment costs. This is because the discontinuation of a treatment

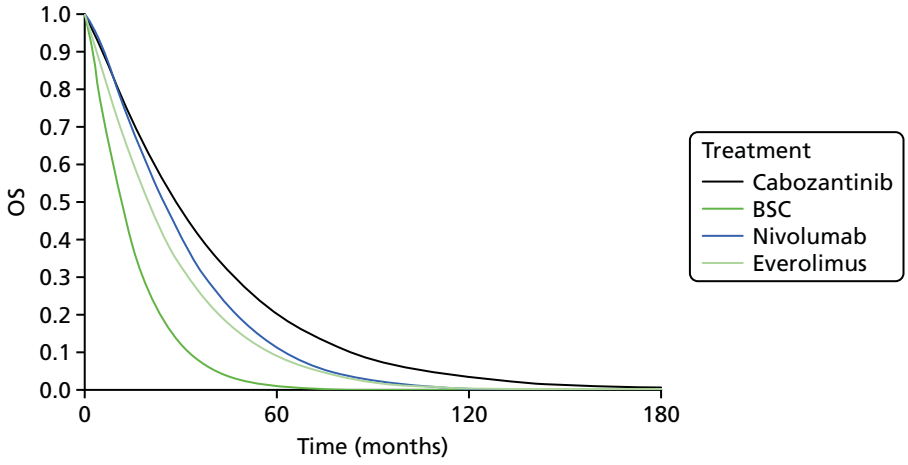


FIGURE 10 Overview of selected OS curves for all treatments.

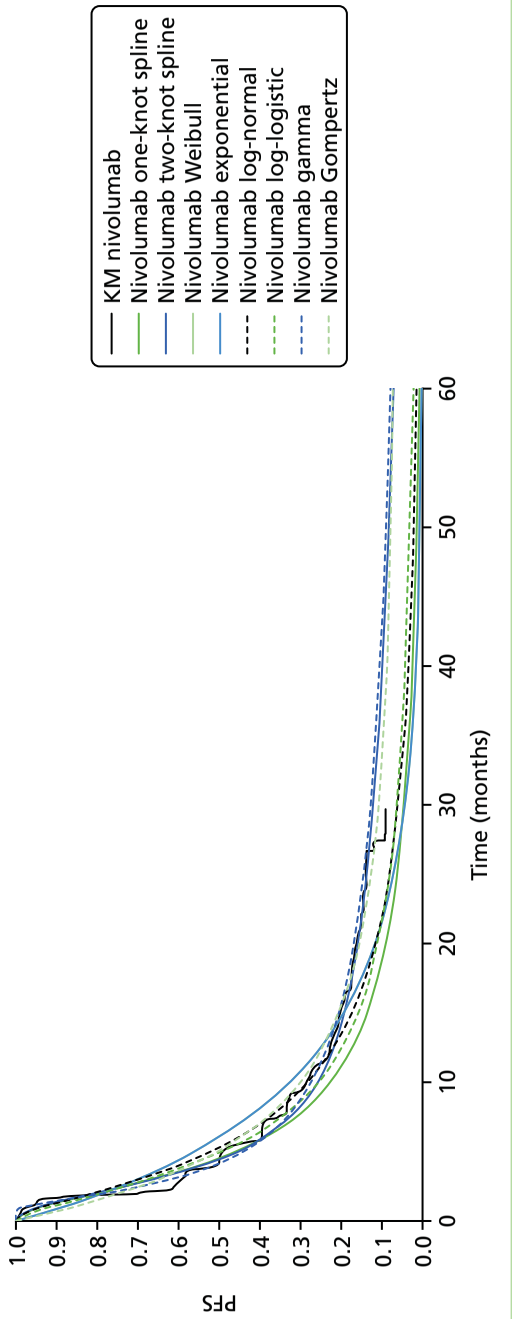


FIGURE 11 Nivolumab KM plot and fitted curves for PFS.

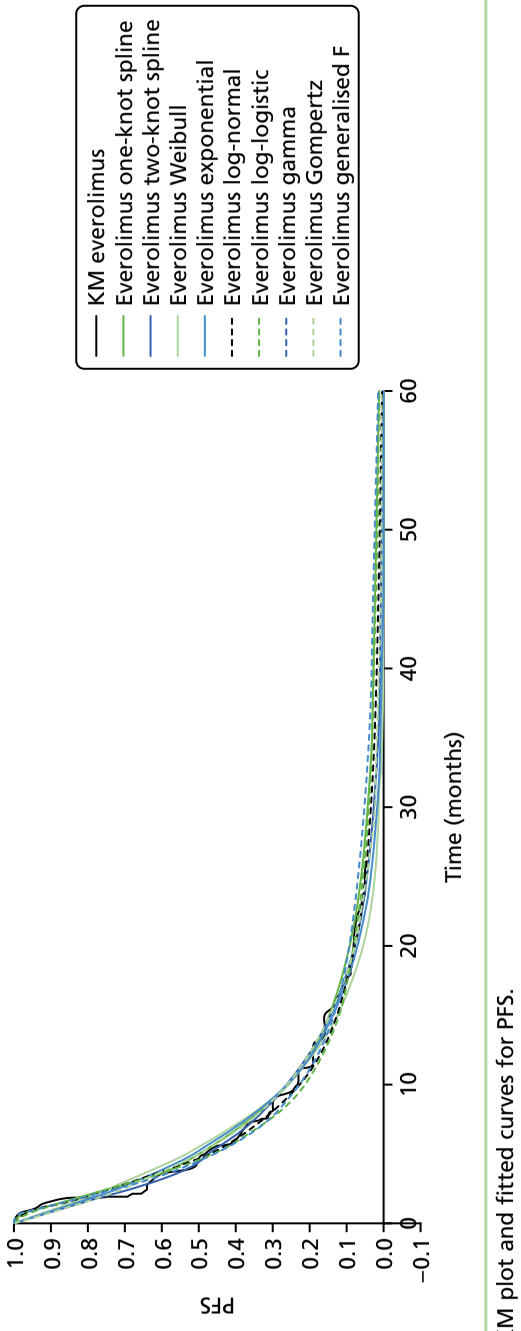


FIGURE 12 Everolimus KM plot and fitted curves for PFS.

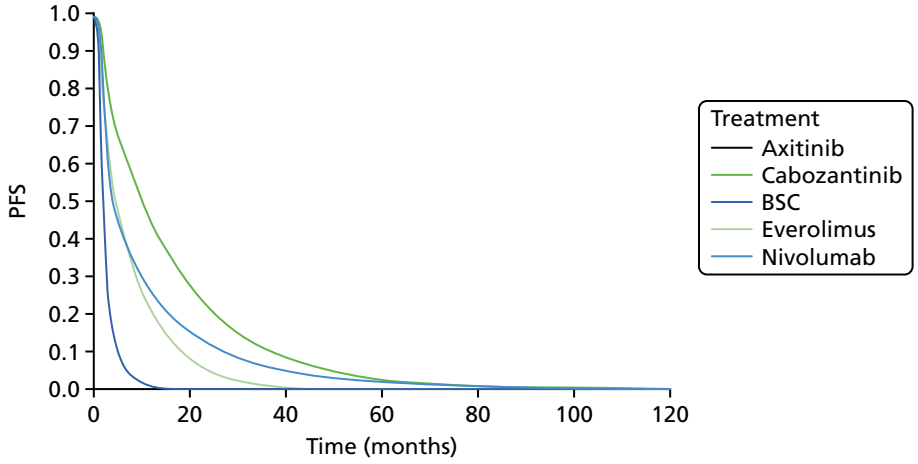


FIGURE 13 Overview of selected PFS curves for all treatments.

(e.g. cabozantinib in the METEOR trial) is, in part, caused by intolerable toxicity caused by the drug, which may not necessarily be reduced if it were given to the population with a different prognosis in another trial (e.g. CheckMate 025). Therefore, the AG chose to fit independent curves to the TTD data to avoid underestimating the treatment costs. This approach may overestimate the treatment costs for cabozantinib if the population characteristics that differ between CheckMate 025 and METEOR are correlated with the time of treatment discontinuation.

The approach taken for PFS and OS was, therefore, also applied to the TTD KM data from the CheckMate 025 trial and the cabozantinib group of the METEOR trial. The regenerated KM plots with fitted curves are given in *Figures 14–16* for everolimus, nivolumab and cabozantinib, respectively. The goodness-of-fit statistics are given in *Appendix 15*.

The two-knot spline had the best fit for the everolimus group with the lowest AIC and BIC, while for nivolumab the two-knot spline had the lowest AIC and was very close to the lowest BIC. For cabozantinib, the two-knot spline did not have as good a fit, with four other models having a better fit based on AIC and BIC. The AG chose to use the two-knot spline for all models to keep the same distribution type across treatment groups and considered this reasonable owing to the relatively small difference in the AIC and BIC statistics for the two-knot spline compared with the best-fitting log-normal for the cabozantinib group. The clear difference in AIC and BIC statistics in favour of the two-knot spline in the everolimus group also supported this decision (curves shown in *Figure 17*). The log-normal distribution was tested as scenario analysis, which is presented in *Results*.

For axitinib, TTD data were not available, so treatment was assumed to discontinue at the point of progression. This is likely to underestimate the treatment cost for axitinib, as the marketing authorisation allows for treatment beyond progression. A scenario analysis assuming that treatment discontinues at the point of progression for all treatments was performed to assess the model results without a discrepancy between treatment assumptions. The results of this are given in *Results*.

Adverse events

The economic model includes grade 3 (or higher) TRAEs. The AG sought clinical expert opinion to understand which TRAEs are expected to have an impact on NHS costs and RCC patients' quality of life, and subsequently included them. However, only TEAEs were identified for cabozantinib and, therefore, included in the model. A summary of the AEs included for each treatment is presented in *Appendix 14*.

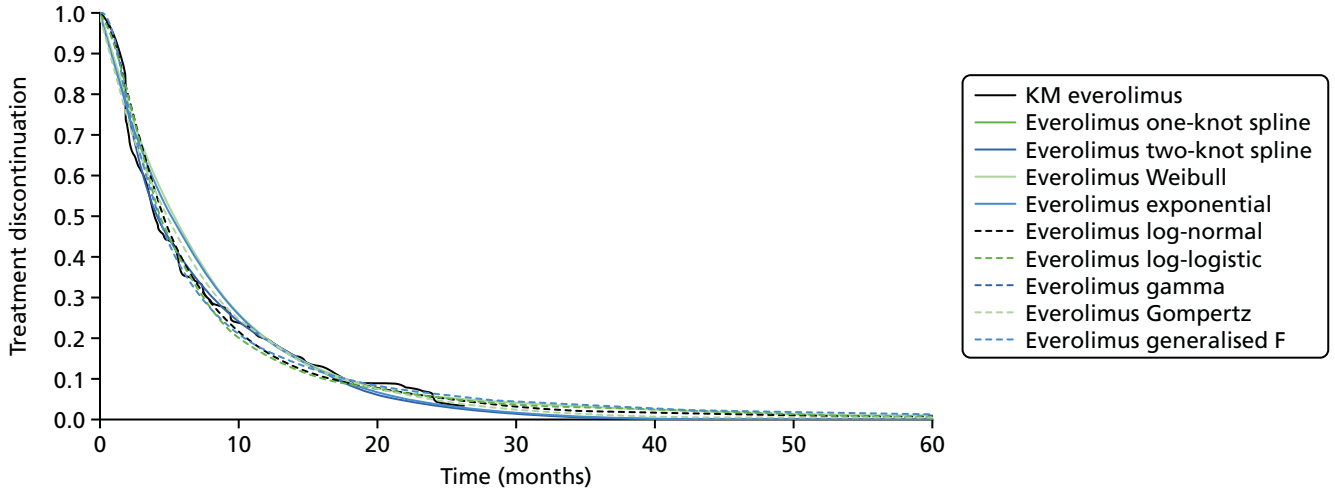


FIGURE 14 Everolimus KM plot and fitted curves for TTD.

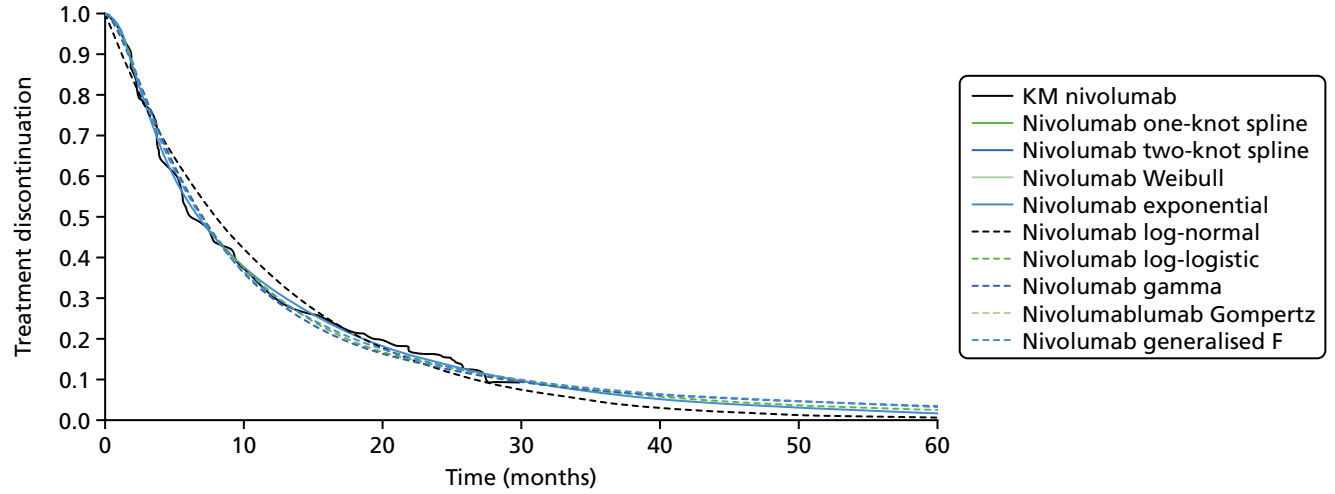


FIGURE 15 Nivolumab KM plot and fitted curves for TTD.

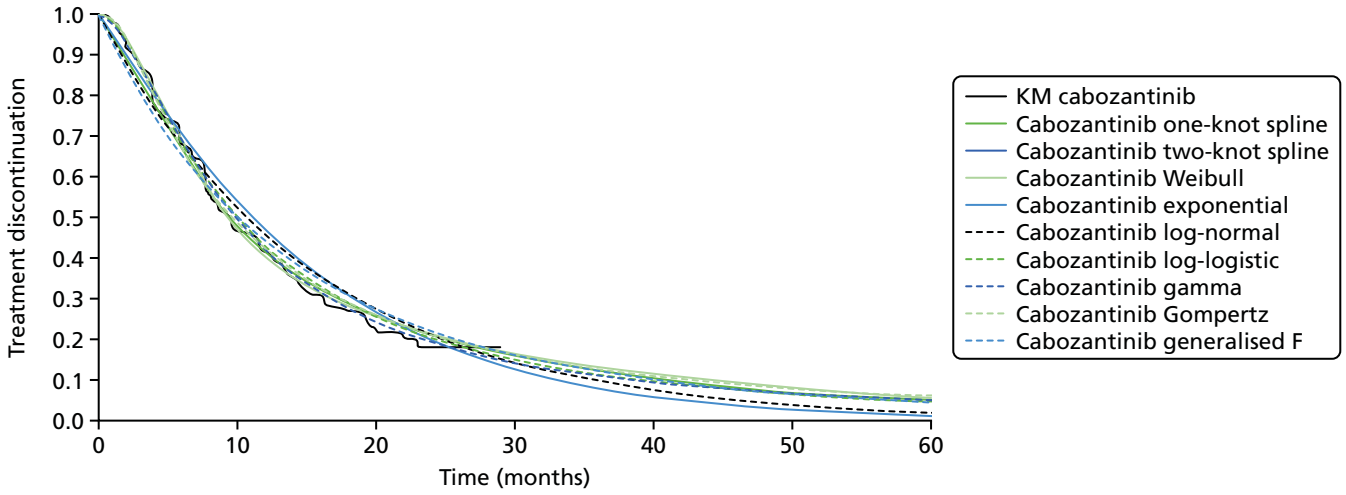


FIGURE 16 Cabozantinib KM plot and fitted curves for TTD.

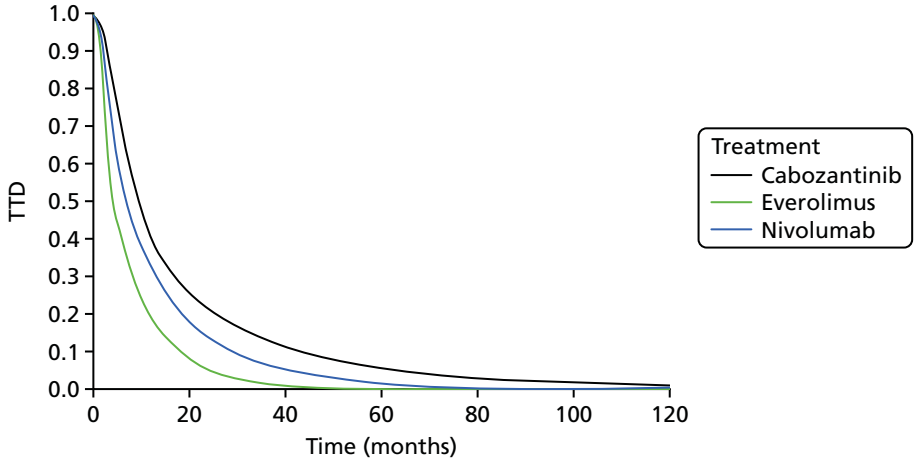


FIGURE 17 Overview of TTD curves for all treatments.

The AG's model captures the impact of TRAEs on costs but assumes that the impact on patients' quality of life is already incorporated into the HSUVs. Therefore, no utility decrements are applied for patients experiencing TRAEs. The AG acknowledges that this is a simplifying assumption, but clinical expert opinion agreed that it is clinically valid (Dr Lisa Pickering, The Royal Marsden NHS Foundation Trust, 2016, personal communication; Professor Martin Gore, The Royal Marsden NHS Foundation Trust, 2016, personal communication; and Dr Amit Bahl, University of Bristol, 2016, personal communication).

Health-related quality of life

Systematic review of existing health-related quality-of-life data

A systematic review was carried out in July 2016 to identify relevant published HRQoL evidence to populate the economic model. The following databases were searched:

- MEDLINE (via Ovid MEDLINE 1946 to present)
- EMBASE (via EMBASE 1974 to week 27 2016)
- Central Register of Controlled Trials (via CENTRAL, The Cochrane Library)
- HTA database (via The Cochrane Library)
- NHS EED (via The Cochrane Library).

The search strategy for all databases combined disease terms and quality of life terms to capture RCC and HRQoL. The search terms that were used are presented in *Appendix 1*. In addition, reference lists of studies identified in the search were scanned to identify potentially relevant papers. No restrictions were applied to any of the searches. Studies were assessed for inclusion based on the criteria outlined in *Appendix 2*. Papers reporting on quality of life of patients with RCC receiving first-line therapy were only included if no relevant studies were identified for patients receiving second-line treatment.

A total of 2200 studies were identified in the database search. The titles and abstracts were assessed by two reviewers. Out of these, 148 were identified as duplicates and 1968 studies were excluded on the basis of title and abstract. A total of 36 studies were identified from the abstract as reporting either condition-specific measures of HRQoL or generic non-preference-based measures of HRQoL. There were 26 studies reporting HRQoL in patients receiving first-line treatment for RCC (23 using generic preference-based measures). Therefore, 59 studies were provisionally included (the full papers of first-line treatment for RCC were only to be reviewed if no studies reporting the use of generic preference-based measures of HRQoL in patients receiving second-line treatment were identified). In cases for which type of measure used or line of treatment was unclear from the abstract, the paper was conservatively considered 'Q1' as described in *Appendix 2* and was ordered for review. The PRISMA diagram for the search is presented in *Figure 18*.

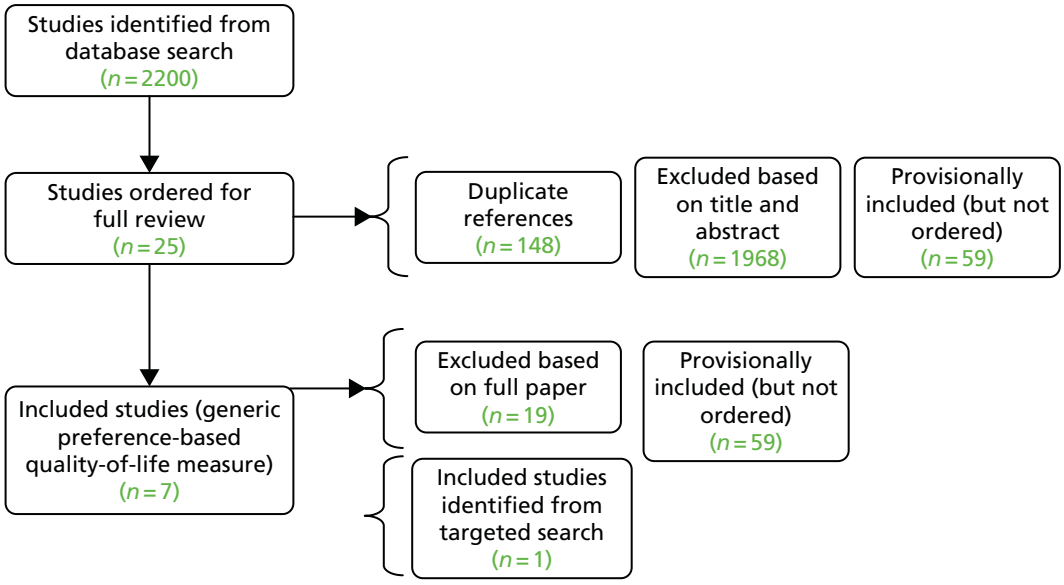


FIGURE 18 The PRISMA flow diagram for HRQoL search.

A total of 25 studies were deemed potentially relevant and were reviewed in full. Out of those, 18 studies were excluded for the following reasons: 'secondary sources of utility values', $n = 5$;[76,78,80,81,84] 'first-line', $n = 4$;[119–122] 'no generic preference-based measures used', $n = 4$;[43,123–125] 'irretrievable', $n = 1$;[126] 'line of treatment not stated', $n = 2$;[127,128] 'insufficient data', $n = 1$;[79] and 'intervention not of interest, $n = 1$.[129]

Six studies were included from the search.[67,82,92,104,130,131] A targeted search aiming to identify additional HRQoL studies in patients receiving nivolumab was carried out after the initial database search. One study was identified and included in this review which was published after the search was carried out.[68] The included studies are described in the following subsection.

Narrative description of published cost-effectiveness studies

Cella and Beaumont[132]

This study assessed the difference in HRQoL between the two treatment arms in CheckMate 025, the open-label Phase III trial comparing nivolumab to everolimus. HRQoL in the trial was measured using FKSI-DRS and EQ-5D questionnaires.

The HRQoL data were collected at baseline for 86% of trial patients, including 362 (88%) of 410 randomised patients in the nivolumab group and 344 (84%) of 411 randomised patients in the everolimus group. Mean EQ-5D scores at baseline were 0.78 (SD 0.24) and 0.78 (SD 0.21) for patients receiving nivolumab and everolimus, respectively. EQ-5D visual analogue scale (VAS) scores were also similar at baseline, with a mean score of 73.3 (SD 18.5) and 72.5 (SD 18.7) for nivolumab and everolimus, respectively.

The EQ-5D utility index and EQ-5D VAS scores improved for patients in the nivolumab arm from baseline to week 104, while scores deteriorated for patients in the everolimus arm. A longitudinal mixed-effects model was used to analyse the HRQoL data. The difference between the two groups (nivolumab compared with everolimus) in EQ-5D utility index was 0.04 (95% CI 0.02 to 0.07; $p = 0.0003$) and in EQ-5D VAS was 5.7 (95% CI 3.8 to 7.7; $p < 0.000$).

The HRQoL analysis was restricted to on-study assessments, which the authors justified by the fact that fewer patients were available to complete the questionnaires at follow-up visits. The authors reported that at the end of treatment, in both groups, most patients who had discontinued treatment had done so because of disease progression.

Cella et al.[133]

The paper reports HRQoL data collected in the Phase III AXIS trial, assessing the effectiveness of axitinib compared with sorafenib in patients who progressed on first-line treatment for RCC.

The number of patients randomised in the trial was 723. Patient-reported outcomes were assessed using the FKSI-15 and the EQ-5D, with questionnaires completed at screening, after every 4 weeks of therapy, at the end of study treatment and at follow-up (28 days after end of therapy). End-of-treatment and follow-up data were collected at different cycles, reflecting the different times that patients went off treatment. Analysis of HRQoL was based on the ITT population and a repeated measures mixed-effects model was used. The model included terms for treatment, time and treatment-by-time with baseline as a covariate and time was assumed to be continuous.

The index scoring algorithm used was that reported by Kind *et al.*,[134] which was derived from a UK general population sample. The mean EQ-5D scores were 0.71 and 0.69 for axitinib and sorafenib, respectively. The corresponding EQ-5D VAS values were 68.11 and 68.64 for axitinib and sorafenib, respectively.

Similar EQ-5D means were observed across the two groups until end of treatment, after which scores declined substantially mainly due to disease progression. There was no statistically significant difference between axitinib and sorafenib in EQ-5D index scores in the mixed-effects model with a difference in means of 0.02 ($p = 0.190$) or in the interaction between treatment and time ($p = 0.8048$).

Cella et al.[130]

The paper reports the analysis of time to deterioration of HRQoL in patients in the double-blind RCT comparing pazopanib with placebo. The total number of patients in the trial was 435 (290 in the pazopanib arm and 145 in the placebo arm), which included 200 patients with prior cytokine therapy and 235 treatment-naive patients. HRQoL estimates were available for 328 patients.

The EQ-5D and EORTC QLQ-C30 questionnaires were used to collect HRQoL data. The data were collected at the following time points: baseline and period of best response (the period between the first date of best response and progression was considered a period of best response). A published preference-based algorithm derived from the UK general population was used to calculate the EQ-5D score.[135] The EQ-5D scores are summarised in *Appendix 13*.

The effect of pazopanib on time to $\geq 20\%$ decline of baseline scores was estimated overall and stratified according to line of treatment. HRQoL was also stratified by benefit (i.e. complete or partial response versus PD, complete or partial response versus stable disease and stable disease versus progressive disease). Patients whose best response was PD experienced greater deterioration of HRQoL than patients whose best response was complete response or stable disease.

Patients who did not experience HRQoL deterioration from baseline score by at least 20% were censored at the time of their last assessment. Time to the first deterioration was estimated using a Cox PH model with the treatment arm as the covariate of interest. Models were estimated with and without controls for baseline HRQoL scores. Analyses were performed on the sample of all patients and on samples of patients stratified by line of therapy (treatment-naive vs. pre-treated with cytokine).

Patients in the pazopanib arm had a lower risk of at least 20% deterioration than those in the placebo arm. However, this difference was not statistically significant across both the treatment-naive and the pre-treated group.

Karakiewicz et al.[131]

This paper reported HRQoL data from a single-arm Phase II study of axitinib. The study was carried out in 15 patients with metastatic renal cell carcinoma (mRCC) whose disease progressed after one prior systemic first-line regimen containing single or combination therapy with a cytokine, VEGF or mTORi, or those who discontinued owing to prohibitive toxicity.

Patient-reported outcomes were assessed using the EQ-5D questionnaire administered on the first day of the first treatment cycle before dosing and before any other clinical assessments, then every 4 weeks while on study, at the end of treatment/withdrawal and, finally, at follow-up (28 days after last dose). The mean EQ-5D score at baseline was 0.7947 and the mean change from baseline to end of treatment was −0.0837. The mean EuroQol VAS score at baseline was 73.3 and the mean change from baseline to end of treatment was −6.5.

Motzer et al.[92]

The paper reports HRQoL data collected in a multicentre, Phase II clinical trial assessing the efficacy of sunitinib monotherapy on cytokine-refractory metastatic RCC in 63 patients.

The EQ-5D questionnaires were filled in by 60 patients at baseline. The mean and median baseline health state EQ-5D VAS scores were 77.1 and 80.0, respectively. The authors reported that the study population's quality of life before sunitinib treatment was similar to that of an age-matched US general population. Mean and median health state VAS scores were similar to the baseline scores throughout the 24 weeks of treatment.

Escudier et al.[107]

This paper reports the analysis of HRQoL data collected in an open-label, multicentre Phase II RCC study assessing the effectiveness of sunitinib when administered on a continuous once-daily dosing regimen. A total of 107 patients were randomly assigned to one of two groups, with 54 patients receiving sunitinib in the morning (a.m. group) and 53 in the evening (p.m. group).

Patient-reported outcomes were assessed using the FKSI-15 and EQ-5D, and were completed at the following time points: screening (baseline measure), after every 4 weeks of therapy, at end of study treatment and finally at follow-up (28 days after end of therapy).

Patients went off treatment at different times and, therefore, end-of-treatment and follow-up data were collected at different cycles to reflect this. Fifty-two patients in the a.m. group and 52 in the p.m. group filled in baseline EQ-5D and Functional Assessment of Chronic Illness Therapy – Fatigue questionnaires, evaluating patient-reported general health status and fatigue, respectively.

The median baseline scores were identical in each treatment arm for the two EQ-5D measurement periods, with utility scores of 0.8 for both the a.m. and p.m. groups, and EQ-VAS scores of 70 for both the a.m. and p.m. groups. The authors reported that EQ-VAS was slightly lower than that of an age-matched sample of the general US population (aged 55–74 years, score of 82), males (score of 83) or respondents with a chronic medical condition (score of 80). The EQ-5D index and EQ-VAS scores did not change from baseline up to 29 cycles of treatment in either cohort and there were no statistically significant differences observed between the cohorts.

Thompson Coon et al.[82]

This paper is part of the published HTA report for the NICE TA178, which is described in detail in *Systematic review of existing cost-effectiveness evidence*. Sunitinib and sorafenib were assessed as second-line treatments for RCC among other first-line treatments. The HSUVs used by the ERG for sunitinib and sorafenib were based on values presented in the company submission for sunitinib as second-line treatment, that were derived from EQ-5D data collected in the single-arm Phase II sunitinib trial. The ERG stated that despite being unable to verify the methods used, the values reported in the company's submission was the best available data in the absence of published data on HSUVs in RCC.

The ERG assumed HSUVs to be the same regardless of treatment received. The HSUVs used in the model were 0.76 and 0.68 for PFS and PD, respectively. Additional sources of HSUVs in RCC were not identified in database searches. Only one source of HSUVs in RCC was identified in the HRQoL database search and that was the paper published by Thompson Coon *et al.*,[82] which did not contain methodological details. Therefore, a previous HTA was reviewed to identify values that could be used in the model that were based on robust methods.[82]

Health state utility values presented in TA333 evidence review group report[89]

As reported in *Single technology appraisal number 333: axitinib for treating renal cell carcinoma after failure of prior systemic treatment*, the EQ-5D analysis in TA333 was based on the full trial population as there was no statistically significant difference between the prior-sunitinib and prior-cytokine subgroups. The estimated mean utility values for PFS and PD were 0.69 and 0.61, respectively. The HSUV for PFS was estimated by calculating the average of the EQ-5D index at every time point in the trial weighted by the number of patients still on treatment. As for PD, the value was based on the weighted average of mean utility at the end of treatment.

Health state utility values presented in TA417 evidence review group report[90]

As reported in *Single technology appraisal number 417: nivolumab for previously treated advanced renal cell carcinoma*, HSUVs for patients receiving nivolumab and everolimus were derived from EQ-5D data obtained from the CheckMate 025 trial. A mixed-effects model was used, with fixed effects for the effects of progression status and treatment allocation; a variable effect for the interaction between treatment arm and progression status; and a random effect for the subject. For patients receiving axitinib and BSC, HSUVs were derived from EQ-5D data from the AXIS trial, reported in TA333. It was assumed that patients who were receiving BSC would experience the same quality of life as patients receiving axitinib. The HSUVs before progression were 0.80, 0.76, 0.69 and 0.69 for nivolumab, everolimus, axitinib and BSC, respectively. The HSUVs after progression were 0.73, 0.70, 0.61 and 0.61 for nivolumab, everolimus, axitinib and BSC, respectively.

Health-related quality-of-life data selected for the economic analysis

In the baseline analysis, the HSUVs reported in TA333 are used for everolimus, axitinib, cabozantinib and BSC. These values were chosen because the methods used in TA333 to estimate HRQoL were clearly reported and were derived from EQ-5D data collected in the AXIS trial whose population reflects RCC patients encountered in UK clinical practice. Therefore, the values used in the model are 0.69 and 0.61 for PFS and PD, respectively.[89]

Patients in the model receiving everolimus, axitinib, cabozantinib and BSC are assumed to experience the same quality of life, which only differed according to progression status. Patients receiving nivolumab are assumed to enjoy a superior quality of life before progression, as the CheckMate 025 trial EQ-5D results show that patients' quality of life continues to improve throughout the trial period while on treatment.[68] These assumptions have also been validated by the AG's clinical experts, who have agreed that patients on immunotherapy do experience a higher quality of life than patients on the other types of RCC treatments. Furthermore, the AG's clinical experts confirmed that patients on BSC may experience a similar quality of life as patients on active treatments because they do not suffer from the AEs associated with drugs. The impact of TRAEs on quality of life is assumed to be captured in the HSUVs used in the model and, therefore, is not modelled separately.

The values used for nivolumab in the economic analysis are 0.73 and 0.65 for PFS and PD, respectively. These values were estimated by applying a quality-of-life increment of 0.036 to the HSUVs for other treatments. The increment used was taken from the Cella *et al.*[68] analysis, which is the difference in EQ-5D utility values according to treatment arm in the CheckMate 025 trial. These values were validated by expert clinical opinion.

In summary, the HSUVs used in the AG's economic model, presented in *Table 14*, were chosen for the following reasons.

- Expert clinical opinion (Dr Lisa Pickering, personal communication; Professor Martin Gore, personal communication; and Dr Amit Bahl, personal communication) considered the AXIS trial population reflective of patients seen in UK clinical practice. Expert clinical opinion sought by the AG (Dr Lisa Pickering, personal communication; Professor Martin Gore, personal communication; and Dr Amit Bahl, personal communication) also indicated that the CheckMate 025 population was less reflective of RCC patients in the UK and could potentially reflect a population with better prognosis, which could bias the quality of life experienced by patients at baseline.
- The HRQoL analysis in TA333 was robust and reported in detail in the ERG report.

Resource use and costs

The following costs were included in the economic model: intervention costs, disease management costs (PFS, PD and TC) and TRAE management costs. All costs were in 2015 Great British pounds.

Drug acquisition and administration costs

The doses considered in the model were axitinib (5 mg, twice daily), cabozantinib (60 mg, once daily) everolimus (10 mg, once daily) and nivolumab (3 mg/kg, every 2 weeks). A summary of the doses and acquisition costs of the treatments included in the model is presented in *Table 15*.

TABLE 14 Health-related quality of life estimates used in the de novo economic analysis

Treatment	State	Utility value
Nivolumab	PFS	0.73
	PD	0.65
Everolimus	PFS	0.69
	PD	0.61
Axitinib	PFS	0.69
	PD	0.61
Cabozantinib	PFS	0.69
	PD	0.61
BSC	PFS	0.69
	PD	0.61

TABLE 15 Intervention costs

Treatment	Formulation[102]	Pack size[102]	Price[102]	Dosage[102]
Axitinib	Inlyta® 5 mg film-coated tablets	56 tablets	NHS indicative price = £3517.00 (Hospital only)	5 mg twice daily
Everolimus	Afinitor® 10-mg tablets	30 tablets	NHS indicative price = £2673.00	10 mg once daily
Nivolumab	Opdivo® 10 mg/ml vials	40-mg vial, 100-mg vial	40 mg, NHS indicative price = £439 100 mg, NHS indicative price = £1097	3 mg/kg every 2 weeks
Cabozantinib	Cabometyx®	20-mg tablets, 40-mg tablets, 60-mg tablets	£5143	60 mg once daily

All of the drugs are administered orally except for nivolumab. The administration cost for nivolumab was assumed to be £185.53 (*NHS Reference Costs 2014–15*,[136] Outpatient, Simple parenteral chemotherapy, Currency code SB12Z).

Disease management costs

There were no UK costing studies for RCC identified in the systematic literature review carried out to identify costing studies. Therefore, the estimates used in the model were based on those used in NICE TA333[89] and TA10078,[137] complemented by expert clinical opinion sought by the AG.

Before progressing, patients are assumed to receive CT scans every 3 months and to be seen by a consultant at the time of the scans, as the scans need clinical revision for signs of progression. Patients also were assumed to have a blood test every month within their progression-free period.

After progressing, patients are assumed to have daily pain medication and 20 community nurse visits per year. A summary of the resource use and costs estimated for each health state is presented in *Table 16*.

Terminal care costs

The estimated cost of TC is £7713. All patients who die in the model incur this cost. The cost estimate is taken from a paper published by the Nuffield Trust estimating costs associated with end-of-life care in the UK. Estimates for patients with a history of cancer were presented separately in the paper.[139] A summary of the cost components of TC used in the paper is presented in *Table 17*.

Adverse event costs

Resource use that was assumed for the management of TRAEs in the model is summarised in *Table 18*. This was validated by the AG's clinical experts.

Subsequent therapy costs

Currently only nivolumab is approved as third-line treatment option for RCC in the UK. However, patients in all the three pivotal trials could receive a variety of further therapies after discontinuing treatment; therefore, the effectiveness estimates from the trials also included the impact of subsequent therapies.

TABLE 16 Resource use and costs for different health states

Health states	Resource	Frequency	Unit cost (£)	Description
PFS	CT scan	Every 3 months	136	*NHS Reference Costs 2014–15*,[136] computerised tomography scan of more than three areas (RD27Z)
	Consultant visit	Every 3 months	189	*NHS Reference Costs 2014–15*,[136] 'Consultant led, first attendance, non-admittance (Code 370 - medical oncology)', WF01B
	Blood test	Every month	3.01	*NHS Reference Costs 2014–15*,[136] directly assessed pathological services – haematology, DAPS05*f*
PD	Specialist community nurse visit	20 visits per year	75	PSSRU,[138] section 10.4 p. 172, Nurse specialist (community), 1-hour patient-related work, including qualifications
	Pain medication	Every day (opioid analgesics: 1 mg/ml vial of morphine sulphate per day)	5.25	BNF[102] morphine sulphate 1 mg/ml

PSSRU, Personal Social Services Research Unit.

TABLE 17 Costs associated with TC

Health states	Resource	Total (in 3 months preceding death)	Unit cost per patient (£)	Description
Terminal care	GP visits	11.4 visits	44	PSSRU,[138] section 10.8 p. 177, GP – unit costs, patient contact lasting 11.7 minutes, including direct staff costs, including qualification costs
	District nurse	7.5 hours	67	PSSRU,[138] section 10.4 p. 172, Nurse specialist (community), 1-hour patient-related work, including qualifications
	Hospital costs	Not available	6239	Emergency, non-emergency, outpatient and A&E visits
	Local authority funded care	Not available	470	Home care, nursing care, residential care, day care, direct payments and respite care
Total cost	£7713			

A&E, accident and emergency; PSSRU, Personal Social Services Research Unit.

TABLE 18 Resource use for management of TRAEs

Serious grade III/IV TRAE	Cost per episode (£)	Source
Hypertension	9	PSSRU,[138] section 10.8, p. 177, GP – unit costs, patient contact lasting 11.7 minutes, including direct staff costs, including qualifications, £44
		PSSRU,[138] section 10.1, p. 169, Community nurse, 1-hour patient time, including qualifications, £67
		5 mg of amilodipine once a day for 4 weeks (cabozantinib STA company submission and BNF) (5 mg, net price 28-tablet pack = 91p)
Pneumonitis	94	One GP visit. PSSRU,[138] section 10.8, p. 177, GP – unit costs, patient contact lasting 11.7 minutes, including direct staff costs, including qualification costs = £44. Average across both arms is 2.93 weeks = £128.92 per episode (PSSRU[138])
		4 weeks of steroids: fluticasone propionate, 50 μg per inhalation, 60 inhalations = £6.38 (based on 100 mg, i.e. two inhalations, per day for 30 days) (MIMS)[140,141]
Diarrhoea (immune-mediated)	192	*NHS Reference Costs 2014–15*,[136] 'Consultant led, first attendance, non-admittance (code 370 - medical oncology)', WF01B = £189
		Loperamide (dose for acute diarrhoea, 4 mg initially, then 2 mg after each loose stool; maximum of 16 mg daily, from BNF,[102] assuming the entire prescription is filled) 2-mg capsule, 30 = £2.98
Anaemia	422	Regular day and night admission SA04J, iron deficiency anaemia with a CC score of 6–9[42]
PPE	101	One GP visit. PSSRU,[138] section 10.8 p. 177, GP – unit costs, patient contact lasting 11.7 minutes, including direct staff costs, including qualification costs:[138] = £44 per episode. Corticosteroid cream (clobetasol) for 50 days

CC, complication or comorbidity; MIMS, Monthly Index of Medical Specialties; PPE, palmar–plantar erythrodysesthesia; PSSRU, Personal Social Services Research Unit.

Patients in the axitinib, everolimus and nivolumab arms of the model receive one further treatment line after discontinuing treatment. The proportions of patients who went on to receive subsequent therapies, and their distribution across various therapies, are based on the numbers reported in the trials as presented in *Table 19*. The proportions for subsequent therapy after nivolumab and everolimus are based on proportions observed in CheckMate 025, reweighted after the removal of bevacizumab and sorafenib, which are currently not treatment options for RCC in the UK at any line. The proportions assumed for after discontinuing axitinib and cabozantinib are based on the AXIS and METEOR trials as reported in the NICE TA463, reweighted after removing sorafenib.

The treatment duration for subsequent therapy is assumed to be 15.82 weeks, which is in line with the median duration of treatment reported in a trial assessing third-line treatment in RCC.[142] The dosage and costs of axitinib and everolimus as third-line treatments are the same as those used to model them as second-line treatments as reported in *Drug acquisition and administration costs*. The dosage and costs of pazopanib and sunitinib are summarised in *Table 20*. The cost of subsequent therapy is applied as a one-off cost as soon as patients discontinue treatment.

Miscellaneous costs

According to the AG's clinical experts, around 10% of patients on BSC receive palliative radiotherapy for bone pain due to metastasis. The outpatient costs of preparation for simple radiotherapy and delivery of a fraction of treatment on a megavoltage machine has been included for patients in the BSC arm of the model, which is in line with the NHS clinical commissioning policy for palliative radiotherapy for bone pain.[143] The costs were estimated using the unbundled Healthcare Resource Group codes SC47Z for planning and SC22Z for delivery, which give a total cost of £419, which is applied to 10% of patients receiving BSC in the model.[136]

Accounting for uncertainty

The impact of parameter uncertainty on model results was investigated in both probabilistic and deterministic analyses.

TABLE 19 Distribution of subsequent therapies received

	Subsequent treatment (%)				
Initial treatment	Axitinib	Everolimus	Sunitinib	Pazopanib	BSC
Cabozantinib	17.00	29.00	5.20	0.00	49.00
Axitinib	0.59	46.38	10.11	10.11	33.30
Everolimus	29.83	0.00	11.05	7.40	51.92
Nivolumab	25.83	27.32	7.26	9.61	29.88
BSC	0.00	0.00	0.00	0.00	100.00

TABLE 20 Subsequent therapy unit costs

Treatment	Formulation[102]	Pack size[102]	Price[102]	Dosage[102]
Sunitinib	SUTENT® 50-mg hard capsules	28 capsules	NHS indicative price = £3138.80	50 mg once daily for 4 consecutive weeks followed by a 2-week rest period (4/2 schedule): a 6-week cycle. Adjusted in steps of 12.5 mg, according to tolerability; usual dose is 25–75 mg daily
Pazobanib	Votrient® 400-mg tablets	30 tablets	£1121.00	800 mg orally once daily

Probabilistic analysis

A PSA was performed to assess the impact of parameter uncertainty on the results of the base-case analysis. This was performed using 10,000 samples from distributions fitted to each uncertain parameter that was included in the PSA. When measures of uncertainty were not available, the SD was set at 25% of the mean value of the parameter to allow sensitivity to be captured for all relevant parameters.

Parameters for costs, resource use and dose adjustments were fitted with gamma distributions to ensure non-negative sampled values. The only costs with a measure of uncertainty were those from *NHS Reference Costs*,[136] which provided a mean value as well as an upper and lower quartile. These data were used to estimate the parameters of a gamma distribution by minimising the sum squared error (SSE) between the predicted quartiles and the reported quartiles. This was performed using the *Solver* tool in Microsoft Excel to vary the beta parameter until the minimum SSE had been achieved [based on the generalised reduced gradient non-linear solving method]. The distribution parameters for all other costs, resource usage and dose adjustment values were calculated directly from the mean and SD (assumed to be 25% of the mean).

Adverse event probabilities and HSUVs were fitted with beta distributions to constrain values between 0 and 1. The distribution parameters were calculated using the numbers of patients experiencing the event and the number of patients not experiencing the event in the relevant trials for each treatment group.

For efficacy of treatments, survival curves for OS, PFS and TTD were simulated in *R* using the *mvrnorm* function from the *MASS* package. This generates a specified number of random samples of the coefficients from the fitted survival models, given the arguments for the vector of coefficients and the covariance matrix of these coefficients. The covariance matrix was derived using the *vcov* function in the standard *stats* package included in *R*. Samples of HRs calculated in the NMA were provided by exporting the CODA (Convergence Diagnosis and Output Analysis) from WinBUGS.

A full summary of the distributions applied to the model parameters is given in *Appendix 16*.

Deterministic analyses

All parameters that varied in the PSA were also assessed with OWSAs using 95% confidence limits from the fitted distributions used for the PSA for the range of values tested.

As HRs for PFS and OS were sampled in the PSA using the CODA output from the NMA, the range of values used in the OWSAs was based on the 95% CrIs derived from the NMA. These are presented in *Table 21*.

TABLE 21 Hazard ratios used in OWSAs relative to everolimus

	Treatment	
OWSA HR ranges	Cabozantinib	BSC
OS		
Base case	0.664	1.901
Lower bound	0.527	0.611
Upper bound	0.823	4.567
PFS		
Base case	0.512	3.058
Lower bound	0.414	2.309
Upper bound	0.627	3.975

Scenario analyses

A number of scenario analyses were performed to assess the impact of different model assumptions on the results. The scenarios tested are described in the following list.

1. HRs for OS and PFS based on the clinical SA where the network was extended to include CheckMate 025, AXIS and observational evidence. The HRs for each treatment (including nivolumab) were applied to the everolimus group curves from CheckMate 025.
2. Applying the Gompertz distributions for OS.
3. Applying the exponential distributions for OS.
4. Applying the one-knot spline curve for OS.
5. Applying the two-knot spline curve for OS.
6. Applying the log-normal distributions for TTD.
7. Setting OS for cabozantinib equal to OS for nivolumab.
8. Setting treatment discontinuation to occur at the point of progression for all treatments.

Results

The base-case model results are presented in *Base-case results*, while results of scenario analyses are given in *Results of scenario analyses*. Results of the deterministic OWSAs are given in *Results of deterministic sensitivity analyses* and PSA results are provided in *Results of the probabilistic sensitivity analysis*.

Base-case results

A summary of disaggregated discounted costs is given in *Table 22*. This shows a breakdown of costs by health state for treatment acquisition, disease management both on and off treatment, subsequent therapy and end of life. A summary of discounted and undiscounted QALYs is given in *Table 23* and *Appendix 17*, respectively, showing total QALYs for health state and treatment discontinuation status.

TABLE 22 Summary of costs (discounted)

	Treatment				
Cost (£) component	Cabozantinib	Axitinib	Everolimus	Nivolumab	BSC
PFS costs					
Treatment acquisition	76,661	29,264	20,216	76,661	0
AEs	691	23	576	240	248
Disease management (on treatment)	1779	927	903	1228	330
Disease management (off treatment)	51	0	24	18	0
Total PFS	79,182	30,214	21,719	78,148	579
PPS costs					
Treatment acquisition	12,315	0	303	12,825	0
AEs	111	0	9	40	0
Disease management (on treatment)	673	0	32	484	0
Disease management (off treatment)	4884	4866	4834	4638	3332
Subsequent therapy	2406	4261	2402	3539	0
End of life	6946	7165	7165	7087	7394
Total PPS	27,334	16,292	14,744	28,613	10,725
Total costs[a]	106,516	46,506	36,463	106,761	11,304

a These values are reported incorrectly in the company submission.

TABLE 23 Summary of QALYs (discounted)

	Treatment				
Health state	Cabozantinib	Axitinib	Everolimus	Nivolumab	BSC
PFS QALYs					
On treatment	0.85	0.44	0.43	0.62	0.16
Off treatment	0.02	0.00	0.01	0.01	0.00
Total PFS	0.87	0.44	0.44	0.63	0.16
PPS QALYs					
On treatment	0.12	0.00	0.01	0.09	0.00
Off treatment	0.87	0.87	0.86	0.88	0.60
Total PPS	0.99	0.87	0.87	0.97	0.60
Total QALYs	1.87	1.31	1.31	1.60	0.75

An incremental analysis of the results, both discounted and undiscounted, is given in *Table 24* and *Appendix 17*, respectively.

Results of scenario analyses

The results of the scenario analyses are shown in *Appendix 18*.

Results of deterministic sensitivity analyses

Deterministic OWSAs were performed for every parameter that was varied in the PSA, using an upper and lower limit for each parameter based on the 95% CI for the distributions fitted to each parameter in the PSA. In addition, the discount rate was tested from 0% to 6%. Given the large number of parameters in the model, the results presented in *Appendix 19* are restricted to those analyses in which either the total costs changed by at least £2000 for at least one treatment group or the total QALYs changed by at least 0.01 for at least one treatment group. In addition, given the number of comparators, the results summary contains the change in costs and QALYs for each treatment group but no ICERs are provided. However, the net monetary benefit (NMB) of each treatment is given also in *Appendix 19* for the OWSAs, for which the ranking of NMB changed from the base-case results. This is given at both the £20,000 and £30,000 per QALY threshold.

Results of the probabilistic sensitivity analysis

The mean results from the 10,000 probabilistic samples is given in *Table 25*. A plot of costs and QALYs for all the samples in each treatment group is given in *Figure 19*.

TABLE 24 Incremental cost-effectiveness results (discounted)

Treatment	Total cost (£)	Total QALYs	Total life-years	ICER (£)
BSC	11,304	0.75	1.25	–
Everolimus	36,463	1.31	2.21	44,965
Axitinib	46,506	1.31	2.21	Dominated by everolimus
Cabozantinib	106,516	1.87	3.18	126,230
Nivolumab	106,761	1.60	2.53	Dominated by cabozantinib

TABLE 25 Mean PSA results

Treatment	Total cost (£)	Total QALYs	Total life-years	ICER (£)
BSC	11,860	0.75	1.25	–
Everolimus	37,393	1.32	2.21	45,450
Axitinib	48,026	1.32	2.21	Dominated by everolimus
Cabozantinib	107,979	1.89	3.22	122,733
Nivolumab	108,353	1.60	2.54	Dominated by cabozantinib

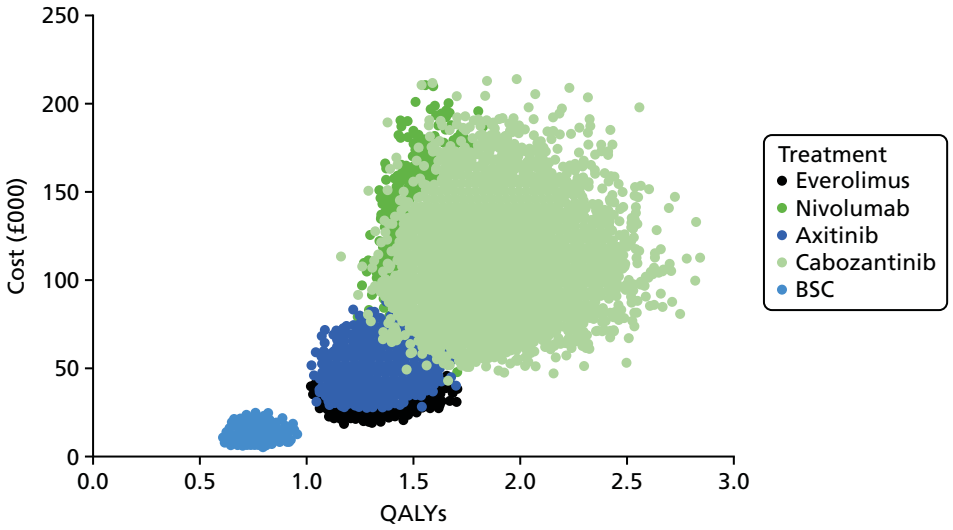


FIGURE 19 The PSA sampled results for all treatments.

Chapter 5 Discussion

Statement of principal findings

Clinical effectiveness

A systematic review was conducted that identified four RCTs (one double-blind and three open label),[43,53,54,57] which met the review inclusion criteria. The paucity of RCT data meant that head-to-head data for all of the interventions and outcomes of interest were not available. In addition, it was not possible to combine all of the RCTs in a network for a MTC owing to the paucity of common comparators across the studies. The inclusion criteria of the review were thus expanded to enable the inclusion of comparative observational studies in an attempt to address the paucity of RCT data for axitinib and sunitinib, which enabled the inclusion of eight non-RCT studies (six retrospective cohort studies and two crossover RCTs for which only the second-phase data were relevant).[55,56,58–63] However, even with the inclusion of non-RCT data, the data available for the interventions of interest were limited and insufficient for addressing all of the outcomes of interest. The non-RCT data were also deemed unsuitable for inclusion in primary analyses due to concerns around study quality and increased risk of bias compared with the RCTs.

The MTCs were conducted as part of the review of clinical effectiveness to enable treatment comparisons to be made between axitinib, BSC (placebo), cabozantinib, everolimus, nivolumab and sunitinib. In the MTC for the primary analysis of PFS, cabozantinib was associated with a statistically significant improvement in PFS compared with everolimus with a HR of 0.51 (95% CrI 0.41 to 0.63). Cabozantinib and everolimus both led to statistically significant benefits in PFS compared with BSC (HR 0.17, 95% CrI 0.12 to 0.24; and HR 0.33, 95% CrI 0.25 to 0.43, respectively). It was not possible to include nivolumab in the analyses of PFS because PHs did not hold for PFS in CheckMate 025,[54] which was the only study evaluating nivolumab included in this review. In addition, axitinib and sunitinib could not be included in the primary analyses for PFS as it was not possible to link AXIS, the only RCT identified for axitinib, into the evidence network and no RCT data were identified for sunitinib. Cabozantinib had a 99% probability of being the most effective treatment for improving PFS.

Sensitivity analyses were conducted for PFS and OS that included non-randomised evidence, in order to include more treatments of interest in the network. Five reasonable quality non-RCTs could be included for PFS[55,56,58,62,63] and, thus, a third RCT;[43] this analysis showed statistically significant benefits for all active treatments over BSC (everolimus HR 0.33, 95% CrI 0.25 to 0.43; cabozantinib HR 0.17, 95% CrI 0.12 to 0.24; axitinib HR 0.31, 95% CrI 0.21 to 0.44; and sunitinib HR 0.27, 95% CrI 0.17 to 0.40). Cabozantinib showed a statistically significant benefit against all other treatments: everolimus (HR 0.51, 95% CrI 0.41 to 0.63), sunitinib (HR 0.63, 95% CrI 0.44 to 0.95), axitinib (HR 0.54, 95% CrI 0.40 to 0.76) and BSC (HR 0.17, 95% CrI 0.12 to 0.24). None of the differences in PFS between sunitinib, everolimus and axitinib was statistically significant.

The MTC for OS included cabozantinib, everolimus, nivolumab, and BSC. Despite PHs not holding for the first 6 weeks in CheckMate 025,[54] the pragmatic decision was made to include this study as 6 weeks is a relatively short amount of time for OS. In addition, RPSFTM crossover-adjusted data were included for RECORD-1 in an attempt to minimise the bias from the BSC patients who crossed over to treatment with everolimus on progression.[53] No data were available to inform OS for sunitinib and so its relative effectiveness in terms of OS is unclear. The results of the MTC for OS suggest both cabozantinib and nivolumab significantly prolong OS compared with everolimus and that there is no significant difference between nivolumab and cabozantinib:

- cabozantinib versus everolimus HR 0.66 (95% CrI 0.53 to 0.82)
- nivolumab versus everolimus HR 0.73 (95% CrI 0.60 to 0.89)
- nivolumab versus cabozantinib HR 1.12 (95% CrI 0.82 to 1.49).

The primary OS analysis suggests that cabozantinib has a 72% probability of being the most effective of the treatments assessed.

The results of the SA for OS, including data to compare treatments with axitinib, were in keeping with those of the primary analysis (no statistically significant benefits of treatments over BSC; nivolumab and cabozantinib benefit over everolimus). Everolimus, cabozantinib and nivolumab all showed longer OS compared with axitinib (HR 0.74, 95% CrI 0.56 to 0.99; HR 0.48, 95% CrI 0.34 to 0.71; and HR 0.54, 95% CrI 0.38 to 0.77; respectively). Data were not available to provide an OS estimate for sunitinib compared with the other treatments. The results of the SA for OS, including data to compare treatments with axitinib, were in keeping with those of the primary analysis (no statistically significant benefits of treatments over BSC; nivolumab and cabozantinib benefit over everolimus). Everolimus, cabozantinib and nivolumab all showed longer OS than axitinib (HR 0.74, 95% CrI 0.56 to 0.99; HR 0.48, 95% CrI 0.34 to 0.71; and HR 0.54, 95% CrI 0.38 to 0.77, respectively). However, it should be noted that there was statistically significant inconsistency in the network for this analysis.

In the network for the SA, axitinib was directly linked to everolimus by one observational study and via sorafenib by one RCT and two observational studies. All three observational studies were rated as being at serious risk of bias, but it is not possible to say how and in what direction this may have affected the results of the studies. Expert clinical opinion also refutes the results of the SA for axitinib. Expert clinical opinion is that axitinib is at least as effective as everolimus, for PFS and OS. This opinion is supported by the PFS results from an indirect comparison of everolimus and axitinib reported by Sherman *et al.*[103] This study conducted a weight-adjusted indirect comparison using sunitinib-refractory subgroup data from two RCTs (AXIS and RECORD-1) that were also included in this review.[43,64] The results from Sherman *et al.*[103] suggest a median PFS of 4.8 months with axitinib (95% CI 4.5 to 6.4 months) and 4.7 months with everolimus (95% CI 3.5 to 10.6 months) with no statistically significant difference between axitinib and everolimus ($p > 0.05$). In addition, the efficacy of axitinib and everolimus were assumed to be equal for both PFS and OS in the recently published NICE guidance for nivolumab in previously treated RCC.[31] This assumption that axitinib and everolimus have equal efficacy for both OS and PFS has been used in the cost-effectiveness analysis of this report as it is deemed to be more clinically plausible than the results generated through the MTC SAs.

Analyses of response rate confirmed the clinical effectiveness of cabozantinib, everolimus and nivolumab compared with BSC and suggests that cabozantinib and nivolumab were the most effective treatments with no statistically significant difference between them (nivolumab vs. cabozantinib OR 1.05, 95% CrI 0.41 to 2.18). The active treatments all resulted in statistically significant improvements in ORR compared with BSC, but small numbers of events led to very wide CrIs:

- everolimus versus BSC (OR 7.14, 95% CrI 1.32 to 8216)
- cabozantinib versus BSC (OR 42.12, 95% CrI 7.55 to 51,921)
- nivolumab versus BSC (OR 41.67, 95% CrI 7.56 to 50,276).

The results of the analyses of PFS and OS suggest a trend towards cabozantinib being the most effective treatment, closely followed by nivolumab with little difference between axitinib, everolimus and sunitinib. All of the active treatments appear to be more effective than BSC.

The HRQoL scores suggested that nivolumab and BSC were favoured over everolimus. Results for everolimus compared with cabozantinib were inconclusive with results from METEOR favouring everolimus over cabozantinib on one measure of disease-specific quality of life.[57] However, HRQoL scores from METEOR were similar on two measures of general HRQoL.[57] No comparative data for axitinib or sunitinib and the interventions under review were identified.

The AE data were inconsistently reported across the studies and are generally inconclusive. The trial level data are suggestive of a worse grade 3/4 AE profile with cabozantinib compared with everolimus and a better grade 3/4 AE profile with nivolumab compared with everolimus. Risk of detection bias in the AE

analyses was high for AXIS, CheckMate 025 and METEOR because safety assessments were done by investigators who were aware of treatment assignment.[43,54,57] However, safety was overseen by a data monitoring committee in CheckMate 025 and METEOR, which may have reduced the risk in those studies.[54,57] The rates of AEs were higher in METEOR than those in CheckMate 025,[54,57] which may be due to the wider definition of AEs in METEOR,[57] including AEs that were not thought to be treatment related, but may equally be indicative of differences in study populations. More patients who received cabozantinib (71.0%) in METEOR had any grade 3 or 4 AEs than those who received everolimus (59.9%),[57] whereas in CheckMate 025 more patients experienced a grade 3 or 4 AE with everolimus (36.5%) compared with nivolumab (18.7%).[54] No summary data were available for everolimus compared with BSC (RECORD-1) or axitinib compared with any of the other treatments under review in this report.[53]

Analysis of the impact of baseline MSKCC prognostic score and number of prior TKI therapies was limited to that from two studies reporting subgroup data for PFS (METEOR and RECORD-1) and two studies for OS (METEOR and CheckMate 025).[53,54,57] The results demonstrated consistent treatment benefit with both cabozantinib and everolimus compared with BSC for PFS irrespective of baseline MSKCC score or number of prior TKI therapies. For OS, the results suggested a trend in favour of cabozantinib and nivolumab over everolimus irrespective of baseline MSKCC score or number of prior TKI therapies. It should be noted that these results are based on small subpopulations of the RCTs and, as a result, should be interpreted with caution as they may be unreliable.

In summary, the evidence base to inform the efficacy of treatments for previously treated amRCC is limited in terms of the number and quality of reporting of studies providing data on the effectiveness of individual interventions. Analyses of PFS and OS suggest a trend towards cabozantinib being the most effective treatment, closely followed by nivolumab with little difference between axitinib, everolimus and sunitinib. All of the active treatments considered in this review appear to be more effective than BSC.

Cost-effectiveness

A key finding seen throughout the analyses is that axitinib is dominated by everolimus owing to the equal effectiveness assumed between them, simplifying a decision analysis between the two treatments to cost minimisation. Everolimus had lower overall costs in all scenarios, resulting in everolimus being dominant. Although everolimus shows a clear benefit in comparison with BSC, the high treatment acquisition costs that make up the majority of the total treatment costs result in a large ICER of £45,000, which is 50% higher than the upper threshold considered for NICE TAs. However, NICE also allow an increased ICER up to £50,000 per QALY for treatments that qualify as end-of-life care. To qualify, the population must have an expected survival of < 2 years and an improved survival with the intervention of at least 3 months. Everolimus may fall into this category based on the expected life-years predicted by the model of < 2 years for BSC and a gain of > 3 months for everolimus. For cabozantinib and nivolumab, the total treatment costs are similar, despite the cost per cycle of nivolumab being higher at £3477 compared with £2400 for cabozantinib. This is mostly a result of the longer treatment duration experienced by patients on cabozantinib in the METEOR trial compared with the treatment duration experienced by patients on nivolumab in the CheckMate 025 trial.[54,57] Note that the similarity in the costs results in a change in ranking between discounted and undiscounted costs for cabozantinib and nivolumab. This is a result of the extended duration of treatment for cabozantinib, which means there is an increase in future costs in comparison with the greater shorter-term costs of nivolumab, leading to a larger decrease when the increased future discount factor is applied.

The benefits shown by cabozantinib in terms of PFS and OS contributed to a gain in QALYs of 0.27 in the base-case analysis compared with nivolumab, and remaining in favour of cabozantinib for all scenario analyses. Given the similarity in costs, some scenario analyses resulted in a change in the order of magnitude of the total costs between nivolumab and cabozantinib; however, the costs always remained similar. This, and the consistent benefit in favour of cabozantinib across all analyses, resulted in nivolumab always being dominated (in some cases extendedly dominated) by cabozantinib. After dominated treatments were excluded, the analyses simplified to a comparison between BSC, everolimus and cabozantinib. All of the

resulting ICERs for these remaining treatments were well above the NICE thresholds, with the exception of the end-of-life threshold, meaning that everolimus may be the most preferable treatment given the end-of-life criteria.

Deterministic OWSAs showed that the most sensitive parameters were the OS HR and the RDIs for all treatments. The upper value of the OS HR resulted in much poorer outcomes for BSC and consequently everolimus became the optimal treatment at the £30,000 per QALY threshold. RDI changes for axitinib and everolimus changed the ranking of NMB for the two treatments, with a high everolimus RDI or a low axitinib RDI resulting in axitinib being more preferable to everolimus but BSC being optimal overall at the £30,000 per QALY threshold. A low everolimus RDI results in everolimus as the optimal treatment at the £30,000 per QALY threshold. Similarly, the lower nivolumab RDI or the upper cabozantinib RDI result in nivolumab being preferable to cabozantinib at the £30,000 per QALY threshold, but did not change the overall outcome of BSC as the optimal treatment. PSA results showed very little difference compared with the deterministic base-case analysis, with an ICER of £45,000 per QALY for everolimus compared with BSC and an ICER of £123,000 per QALY for cabozantinib compared with everolimus.

The applicability of these analyses to a NHS setting needs to be considered further as they are based on the list prices of the active treatments and do not take into account any PASs, which are commercially confidential. The AG is aware that all treatments compared in this HTA have an agreed PAS in place to provide these treatments on the NHS at a lower cost and, therefore, any conclusions drawn from these analyses are very limited in their applicability to health-care decision-making for the NHS. If we were to assume that the PAS discounts are similar across the treatments, we could surmise that the analysis presented indicates the ranking of treatments as we would expect with similar PASs. This would be a reasonable conclusion, given that the acquisitions costs make up an equally large proportion of the total cost for nivolumab and cabozantinib (84% for each), while the proportion for axitinib and everolimus is also fairly similar between the two (63% and 56%, respectively). This would mean that, under the assumption of similar PASs, cabozantinib is likely to remain dominant over nivolumab and, similarly, everolimus is likely to remain dominant over axitinib. Therefore, the analysis may still simplify to a comparison between cabozantinib, everolimus and BSC. Given that everolimus has been approved by NICE, it may be reasonable to assume that the PAS discount is at least enough to reduce the ICER between everolimus and BSC below £30,000. This would require a discount of around 40% based on this analysis. A similar discount for cabozantinib would result in a much higher ICER of around £76,000 in comparison with everolimus and, therefore, under these assumptions, would not represent value for money.

As a large aspect of the analysis relates to producing survival curves for PFS and OS that fit the trial data well and provide a plausible extrapolation beyond the trial period, the AG considered it important to assess a range of different models to find suitably fitting curves for each treatment group. For this reason, the AG considered flexible spline models in addition to standard parametric distributions, which proved important for PFS, for which the flexibility of the shape of the two-knot spline provided significantly better fitting curves than the standard parametric curves. Although this was not necessary for OS, the independently fitted Weibull curves that were chosen provided different shapes for the nivolumab and everolimus. This highlights a potentially important limitation of the analysis: the survival curves generated for cabozantinib, axitinib and BSC were dependent on the everolimus curve and, therefore, followed the same shape. Hence, the nivolumab curve has a distinctly different shape to the curves for the other treatments and this restriction may have a significant impact on the extrapolation of the curves in comparison with a more flexible approach to incorporate the fitting of survival curves into a MTC. A method used in the cabozantinib TA (TA463) by Ouwens *et al.*[144] attempted this but the ERG found that the method had limitations that resulted in poor fits to the trial data.[26,144] Given that an assumption of PHs was found to be reasonable in METEOR and RECORD-1, the AG consider the methods used for this HTA to be more appropriate. For the scenario analysis using HRs derived from the MTC in which the CheckMate 025 trial and the available observational evidence was included, the survival curves, including the nivolumab curve, are dependent on the everolimus curve, so this analysis uses curves with the same shape as the everolimus curve for all treatments. Therefore, this analysis allows consistency across treatments for this issue.

Another limitation of the analysis is that the two-knot spline used for the TTD curves does not provide the best fit to the cabozantinib group and results in a curve that is consistently higher than the nivolumab curve, whereas the KM plots for TTD showed that the cabozantinib curve appeared to converge towards the nivolumab curve. Therefore, this may overestimate the costs incurred in the cabozantinib group and as a result overestimate the ICER. However, the scenario analysis using the log-normal curve appeared to model this aspect better, albeit with a reduced goodness-of-fit to the nivolumab and everolimus curves, and this had very little impact on the total costs and the resulting ICER. In addition to this limitation, the costs of axitinib are likely to be underestimated owing to the assumption that it was used until progression, whereas in practice it could be used beyond progression. To account for this difference in costing approaches, a progression-based treatment schedule was applied in a scenario analysis, showing an unchanged ICER of £45,000 for everolimus but a much reduced ICER of £108,000 for cabozantinib. However, a further analysis could have been performed to determine whether or not there is a link between PFS and TTD for cabozantinib. This could then link to the adjusted cabozantinib PFS curve rather than assuming that TTD does not need to be adjusted between trials, even though the everolimus groups had different TTD in the CheckMate 025 and METEOR trials.[54,145]

On the whole, the analyses presented have accounted for the limitations, when possible, and have used a range of modelling options to find the most plausible inputs to the model. Expert clinical opinion was sought to validate the inputs to provide a model that most reflected UK clinical practice. The key limitation is that the treatment acquisition costs do not reflect those that apply to the current NHS setting owing to confidentially agreed PAS discounts, but this is a limitation that could not be avoided.

Strengths and limitations of the assessment

Clinical effectiveness

This review was conducted according to methods that were prespecified in a prospectively registered protocol.[42] When changes were made to the methods, primarily on the recommendation of clinical experts, these have been made transparent throughout the report. Study inclusion, data extraction and quality assessment were conducted independently by two or more experienced systematic reviewers to ensure all relevant evidence was included, and to reduce bias and error. Searches were designed to identify unpublished data (conference abstracts) and ongoing studies (trial registries) and all results were checked for analysis and transcription errors. However, studies for cabozantinib, which was added as an intervention of interest to the review after the electronic database searches were run, were identified by clinical experts and the company submission for the NICE STA. As no systematic search for cabozantinib studies was carried out, potentially relevant studies may have been missed, but the risk is deemed to be low.

The primary analyses bring together high-quality evidence from RCTs for the most pertinent outcomes in this population (OS and PFS). When possible, the review conclusions are informed by MTCs to estimate relative treatment effects in the absence of head-to-head evidence. Inclusion criteria were widened to incorporate comparative observational evidence in SAs to substantiate the primary results and to provide estimates for all treatments of interest.

The inclusion of recently approved therapies increases the relevance and timeliness of the review; however, evidence for emerging therapies is often limited to a regulatory trial. Furthermore, the comparator used may not always be the most relevant to UK practice. The small number of trials increased the uncertainty in the analyses in a number of ways. First, the PHs assumption did not hold for PFS in the one trial of nivolumab[54] and so we were unable to estimate nivolumab PFS compared with the other treatments via MTC. Second, relevant RCT data for axitinib are limited to a subgroup analysis conducted in one study that did not connect to the network of other RCTs.[43] Third, imprecision surrounding BSC (informed by one study) led to counter-intuitive non-significant differences compared with BSC.[64] The protocol change to conduct SAs incorporating non-randomised evidence provided relative effects for all treatments of interest;

however, the evidence base underpinning the MTC is less reliable and the uncertainty associated with BSC remained because no further BSC comparisons could be incorporated.

Planned subgroup analyses help to unpick the effect of prior therapies and baseline prognostic score, but these are limited to studies reporting disaggregated data and, thus, do not provide results for all treatments. Other key baseline and study design variables are presented in *Table 5* and summarised in *Model structure*, but there were too few studies informing the MTC to support additional analyses to explore whether or not observed inconsistencies could be explained by design or between-group baseline differences (both within studies and between studies of the same treatment). For example, the higher proportion of patients experiencing AEs on everolimus in METEOR (59.9%) than RECORD-1 (36.5%) might be explained by differences in the way that AEs were defined and recorded, or by differences in the severity of the populations (i.e. permitted prior therapies or MSKCC distribution). Additionally, the extent to which the adjusted analysis controlled for placebo (BSC) to everolimus crossover in RECORD-1 cannot be quantified; this may inflate the effectiveness of BSC compared with everolimus.[64]

Cost-effectiveness

The cost-effectiveness analysis was developed after reviewing previously published economic evaluations, as well as NICE TAs published on the NICE website. The strengths and weaknesses of these analyses were considered before developing the methods for the economic analysis presented in this report.

A range of distributions was tested to fit survival curves to PFS and OS data, including flexible hazard-based spline models. To identify the most plausible curves, the statistical fit of these distributions was assessed and expert clinical opinion was sought to validate long-term extrapolations. Expert opinion was also sought to inform and validate decisions around resource use and quality-of-life assumptions. Therefore, the model is considered to be a good reflection of clinical practice in the UK.

The key limitation of the analysis is the absence of the true drug prices on the NHS, which was a limitation that could not be avoided owing to the confidentiality of discounts given to the NHS. This limits the reliability of the conclusions, which could differ markedly if the discounts are very different for each of the drugs in the analysis.

A large range of SAs and scenario analyses were performed to test the robustness of the model. This included a probabilistic analysis using 10,000 samples of all suitable parameters, including survival curves, which proved that the model was very robust to changes in the parameters.

Chapter 6 Conclusion

Implications for service provision

The evidence base to inform the efficacy of treatments for previously treated amRCC is limited in terms of the number and quality of reporting of studies providing clinical effectiveness data for axitinib, cabozantinib, everolimus, nivolumab, BSC and sunitinib. Analyses of PFS and OS suggest a trend towards cabozantinib being the most effective treatment, closely followed by nivolumab with little difference between axitinib, everolimus and sunitinib. All of the active treatments considered in this review appear to be more effective than BSC. Cabozantinib is not yet available for use on the NHS in England, but it is currently undergoing appraisal by NICE.

The results of the cost-effectiveness analysis may not be reliable owing to the inability to apply confidential PAS discounts agreed between the holder of the marketing authorisation and the Department of Health. Therefore, the costs do not fully reflect those incurred by the NHS and the conclusions may differ if these discounts are significantly different across the different treatments.

Suggested research priorities

The searches of trial registries for new or ongoing studies in amRCC did not identify any studies of potential relevance to this review. However, high-quality RCT data comparing all the available RCC treatment options are required to enable more robust estimates of efficacy of the newer RCC therapies with older treatments. In particular, RCT data for sunitinib and axitinib are required to fully assess how they compare to everolimus and confirm the assumptions made in the cost-effectiveness model relating to the efficacy of axitinib and everolimus.

The HRQoL and AEs data also need to be collected in RCTs in a more standardised approach to enable a direct comparison of the RCC treatments. In particular, HRQoL data are required from standardised measurement tools and questionnaires such as EQ-5D. AEs data are required in both aggregate form, with total number of events, and for select AEs relating to the drugs under investigation, to enable them to be analysed in future meta-analyses.

Owing to the lack of RCT data identified for inclusion in this review, there were limited data for analysis on response rates. This is an important outcome to RCC patients and so clarification on the difference in treatment effects in terms of response rates would be a further area for future research to focus on. This should ideally be from RCTs and would require the use of standardised response categories if they are to be combined in meta-analyses.

The applicability of these findings to the NHS is somewhat limited because existing confidential PASs could not be used in the analysis. Future work using the discounted prices, at which these drugs are provided to the NHS, would better inform estimates of their relative cost-effectiveness.

Acknowledgements

The ERG would like to thank Clare Fiatikoski [*British Medical Journal* (BMJ) Evidence Team] and George Osei-Assibey [BMJ-Technology Assessment Group (TAG)] for their contribution to the literature searching and data abstraction; and Dr Amit Bahl (Consultant Oncologist, Bristol Cancer Institute), Dr Lisa Pickering (Consultant Medical Oncologist, St George's Healthcare Trust), Dr Sarah Rudman (Consultant Medical Oncologist, Guy's Hospital), Dr Penny Kechagioglou (Consultant Oncologist, University Hospitals Coventry and Warwickshire NHS Trust), Professor Martin Gore (Consultant Medical Oncologist, The Royal Marsden NHS Foundation Trust) and Dr Naveed Sarwar (Consultant Medical Oncologist, Charing Cross Hospital) for providing clinical advice throughout the project and for providing feedback on the clinical sections of the report.

The ERG would also like to thank Tracey Jhita (BMJ-TAG), Andrea Berardi (BMJ-TAG), Gianluca Baio (University College London) and William Browne (University College London) for providing feedback on the health economic sections of the report.

Contributions of authors

Steven J Edwards was the overall project lead, supervised the production of the protocol and final report, acted as methodological advisor and acted as guarantor of the report.

Victoria Wakefield provided project management, devised and carried out literature searches for the systematic review of clinical effectiveness, assessed full publications for inclusion, contributed to data extraction and validation, carried out and validated meta-analysis, wrote the sections of the report relating to clinical effectiveness and contributed to the editing of the report.

Peter Cain provided project management and contributed to the development of the economic model, the survival analysis performed and to the appraisal of abstracts retrieved from the economic literature search. He assessed full publications for inclusion and contributed to data extraction and validation, the writing sections of the report relating to cost-effectiveness, and to the editing of the report.

Charlotta Karner contributed to the appraisal of abstracts retrieved from the literature search, assessed full publications for inclusion, contributed to data extraction and validation, carried out and validated meta-analysis, and contributed to the editing of the report.

Kayleigh Kew contributed to data extraction and validation, the writing of the sections of the report relating to clinical effectiveness and to the editing of the report.

Mariana Bacelar contributed to the development of the economic model, the survival analysis performed, the digitisation of KM plots and the appraisal of abstracts retrieved from the economic literature search. She also assessed full publications for inclusion and contributed to data extraction and validation, the writing of the sections of the report relating to cost-effectiveness and the editing of the report.

Natalie Masento contributed to the appraisal of abstracts retrieved from the literature search, assessed full publications for inclusion and contributed to data extraction and validation, the writing of the sections of the report relating to clinical effectiveness and the editing of the report.

Fatima Salih contributed to the appraisal of abstracts retrieved from the economic literature search, assessed full publications for inclusion and contributed to data extraction and validation, the writing of the sections of the report relating to cost-effectiveness, the editing of the report and the digitisation of KM plots.

All authors read and commented on draft versions of the report.

Data sharing statement

This is a systematic review and, therefore, the data used for each analysis are present within the report. Further information and requests for access to the data can be obtained from the corresponding author.

References

1. Cancer Research UK. *Kidney Cancer Statistics*. Oxford: Cancer Research UK; 2015. URL: www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/kidney-cancer#heading-Zero (accessed 10 March 2016).
2. Smittenaar CR, Petersen KA, Stewart K, Moitt N. Cancer incidence and mortality projections in the UK until 2035. *Br J Cancer* 2016;**115**:1147–55. https://doi.org/10.1038/bjc.2016.304
3. Cancer Research UK. *Kidney Cancer Incidence Statistics*. Oxford: Cancer Research UK; 2016. URL: www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/kidney-cancer/incidence (accessed 7 November 2016).
4. Cancer Research UK. *Types of Kidney Cancer*. Oxford: Cancer Research UK; 2016. URL: www.cancerresearchuk.org/about-cancer/type/kidney-cancer/about/types-of-kidney-cancer (accessed 2 November 2016).
5. Ljungberg B, Campbell SC, Choi HY, Jacqmin D, Lee JE, Weikert S, *et al.* The epidemiology of renal cell carcinoma. *Eur Urol* 2011;**60**:615–21. https://doi.org/10.1016/j.eururo.2011.06.049
6. Cancer Research UK. *Stages of Kidney Cancer*. Oxford: Cancer Research UK; 2016. URL: www.cancerresearchuk.org/about-cancer/type/kidney-cancer/treatment/stages-of-kidney-cancer (accessed 7 November 2016).
7. Cancer Research UK. *Kidney Cancer Survival Statistics: Kidney Cancer Survival by Stage at Diagnosis*. Oxford: Cancer Research UK; 2016. URL: www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/kidney-cancer/survival#heading-Three (accessed 6 December 2016).
8. Deng FM, Melamed J. Histologic variants of renal cell carcinoma: does tumor type influence outcome? *Urol Clin North Am* 2012;**39**:119–32, v. https://doi.org/10.1016/j.ucl.2012.02.001
9. Molina AM, Tickoo SK, Ishill N, Trinos MJ, Schwartz LH, Patil S, *et al.* Sarcomatoid-variant renal cell carcinoma: treatment outcome and survival in advanced disease. *Am J Clin Oncol* 2011;**34**:454–9. https://doi.org/10.1097/COC.0b013e3181f47aa4
10. Shuch B, Bratslavsky G, Linehan WM, Srinivasan R. Sarcomatoid renal cell carcinoma: a comprehensive review of the biology and current treatment strategies. *Oncologist* 2012;**17**:46–54. https://doi.org/10.1634/theoncologist.2011-0227
11. Escudier B, Porta C, Schmidinger M, Rioux-Leclercq N, Bex A, Khoo V, *et al.* Renal cell carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2016;**27**(Suppl. 5):v58–68. https://doi.org/10.1093/annonc/mdw328
12. Gupta K, Miller JD, Li JZ, Russell MW, Charbonneau C. Epidemiologic and socioeconomic burden of metastatic renal cell carcinoma (mRCC): a literature review. *Cancer Treat Rev* 2008;**34**:193–205. https://doi.org/10.1016/j.ctrv.2007.12.001
13. Kidney Research UK. *Kidney Cancer*. URL: http://www.kidneyresearchuk.org/health-information/kidney-cancer (accessed 31 January 2017).
14. Heng DY, Xie W, Regan MM, Warren MA, Golshayan AR, Sahi C, *et al.* Prognostic factors for overall survival in patients with metastatic renal cell carcinoma treated with vascular endothelial growth factor-targeted agents: results from a large, multicenter study. *J Clin Oncol* 2009;**27**:5794–9. https://doi.org/10.1200/JCO.2008.21.4809

15. Motzer RJ, Mazumdar M, Bacik J, Berg W, Amsterdam A, Ferrara J. Survival and prognostic stratification of 670 patients with advanced renal cell carcinoma. *J Clin Oncol* 1999;**17**:2530–40. https://doi.org/10.1200/JCO.1999.17.8.2530

16. Karnofsky DA, Burchenal JH. The Clinical Evaluation of Chemotherapeutic Agents in Cancer. In MacLeod CM, editor. *Evaluation of Chemotherapeutic Agents*. New York, NY: Columbia University Press; 1949. p. 196.

17. Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, McFadden ET, Carbone PP. Toxicity and response criteria of the Eastern Cooperative Oncology Group. *Am J Clin Oncol* 1982;**5**:649–55. https://doi.org/10.1097/00000421-198212000-00014

18. National Cancer Intelligence Network. *Kidney Cancer: Incidence, Mortality and Survival Rates in the United Kingdom*. URL: www.ncin.org.uk/view?rid=1002; 2013 (accessed 31 January 2017).

19. Department of Health. *2010 to 2015 Government Policy: Cancer Research and Treatment*. London: Department of Health; 2015. URL: www.gov.uk/government/publications/2010-to-2015-government-policy-cancer-research-and-treatment/ (accessed 1 February 2017).

20. Institute for Health Metrics and Evaluation (IHME). *Global Burden of Disease Compare Data Visualization*. Seattle, WA: IHME, University of Washington; 2016.

21. NHS Choices. *Kidney Cancer – Treatment*. URL: www.nhs.uk/Conditions/Cancer-of-the-kidney/Pages/Treatmentaspx (accessed 17 December 2015).

22. National Institute for Health and Care Excellence. *Improving Supportive and Palliative Care for Adults with Cancer*. London: NICE; 2004. URL: www.nice.org.uk/guidance/csg4/resources/improving-supportive-and-palliative-care-for-adults-with-cancer-773375005 (accessed 1 February 2017).

23. National Institute for Health and Care Excellence. *Pazopanib for the First-line Treatment of Advanced Renal Cell Carcinoma. Technology Appraisal Guidance*. London: NICE; 2011. URL: www.nice.org.uk/guidance/ta215/resources/pazopanib-for-the-firstline-treatment-of-advanced-renal-cell-carcinoma-82600251340741 (accessed 17 December 2015).

24. National Institute for Health and Care Excellence. *NICE Technology Appraisal Guidance TA169. Sunitinib for the First-line Treatment of Advanced and/or Metastatic Renal Cell Carcinoma*. London: NICE; 2009. URL: www.nice.org.uk/guidance/ta169/resources/sunitinib-for-the-firstline-treatment-of-advanced-andor-metastatic-renal-cell-carcinoma-82598383607749 (accessed 21 February 2017).

25. National Institute for Health and Care Excellence. *Axitinib for Treating Advanced Renal Cell Carcinoma After Failure of Prior Systemic Treatment. Technology Appraisal Guidance*. London: NICE; 2015. URL: www.nice.org.uk/guidance/ta333/resources/axitinib-for-treating-advanced-renal-cell-carcinoma-after-failure-of-prior-systemic-treatment-82602545696197 (accessed 2 December 2015).

26. National Institute for Health and Care Excellence. *CDF Rapid Reconsideration: Everolimus for the Second-line Treatment of Metastatic Renal Cell Carcinoma (Review of TA219) [ID1015]*. London: NICE; 2017. URL: www.nice.org.uk/guidance/TA432/documents/committee-papers-5 (accessed 8 March 2017).

27. National Institute for Health and Care Excellence. *Nivolumab for Previously Treated Advanced or Metatstatic Renal Cell Carcinoma – Final Scope*. London: NICE; 2016. URL: www.nice.org.uk/guidance/ta417 (accessed 3 April 2017).

28. National Institute for Health and Care Excellence. *Single Technology Appraisal: Cabozantinib for Previously Treated Advanced Renal Cell Carcinoma [ID931]. Committee Papers*. London: NICE; 2017. URL: www.nice.org.uk/guidance/GID-TA10075/documents/committee-papers (accessed 7 March 2017).

29. National Institute for Health and Care Excellence. *NICE Technology Appraisal Guidance TA178. Bevacizumab (First-line), Sorafenib (First- and Second-line), Sunitinib (Second-line) and Temsirolimus (First-line) for the Treatment of Advanced and/or Metastatic Renal Cell Carcinoma*. London: NICE; 2009. URL: www.nice.org.uk/guidance/ta178 (accessed 21 February 2017).

30. Wagstaff J, Jones R, Hawkins R, Porfiri E, Pickering L, Bahl A, *et al.* Treatment patterns and clinical outcomes in patients with renal cell carcinoma in the UK: insights from the RECCORD registry. *Ann Oncol* 2016;**27**:159–65. https://doi.org/10.1093/annonc/mdv504

31. National Institute for Health and Care Excellence. *Final Appraisal Determination: Nivolumab for Previously Treated Advanced Renal Cell Carcinoma (TA417)*. London: NICE; 2016. URL: www.nice.org.uk/guidance/TA417/documents/final-appraisal-determination-document (accessed 7 March 2017).

32. National Institute for Health and Care Excellence. *Renal Cancer NICE Pathway*. London: NICE; 2017. URL: https://pathways.nice.org.uk/pathways/renal-cancer (accessed 10 March 2017).

33. Electronic Medicines Compendium. *Inlyta 1 mg 3mg, 5 mg and 7mg Film-coated Tablets*. June 2016. URL: www.medicines.org.uk/emc/medicine/27051 (accessed 17 February 2017).

34. Electronic Medicines Compendium. *Cabometyx 20mg, 40mg, 60mg*. October 2016. URL: www.medicines.org.uk/emc/medicine/32431 (accessed 8 March 2017).

35. Electronic Medicines Compendium. *SUTENT 12.5mg, 25mg, 37.5mg and 50mg Hard Capsules*. November 2016. URL: www.medicines.org.uk/emc/medicine/18531 (accessed 17 February 2017).

36. BMJ Technology Assessment Group. *Final Protocol: Axitinib, Everolimus, Sorafenib and Sunitinib for Treated Advanced or Metastatic Renal Cell Carcinoma*. 2015. URL: www.nice.org.uk/guidance/indevelopment/gid-tag526/documents (accessed 7 March 2017).

37. National Institute for Health and Care Excellence. *Single Technology Appraisal: Cabozantinib for Previously Treated Advanced Renal Cell Carcinoma [ID931] – Committee Papers*. London: NICE; 2017. URL: www.nice.org.uk/guidance/GID-TA10075/documents/committee-papers (accessed 7 March 2017).

38. Ispen Limited. *Summary of Product Characteristics (SPC): Cabometyx 20mg, 40mg, 60mg*. 2016. URL: www.medicines.org.uk/emc/medicine/32431 (accessed 7 March 2017).

39. Bayer plc. *Summary of Product Characteristics (SPC): Nexavar 200mg Film-coated Tablets*. 2016. URL: www.medicines.org.uk/emc/medicine/18520 (accessed 7 March 2017).

40. Centre for Reviews and Dissemination. *CRD's Guidance for Undertaking Reviews in Healthcare*. 2009. URL: www.york.ac.uk/media/crd/Systematic_Reviews.pdf (accessed 7 February 2017).

41. Higgins JP, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*. Oxford: Wiley Online Library; 2008. https://doi.org/10.1002/9780470712184

42. Wakefield V, Edwards S, Karner C, Osei-Assibey G, Bacelar M, Berardi A, *et al. Evaluation of the Clinical and Cost-Effectiveness of Everolimus, Nivolumab, Axitinib, Sorafenib and Sunitinib in Previously Treated Renal Cell Carcinoma*. PROSPERO 2016:CRD42016042384. URL: www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42016042384 (accessed 3 April 2017).

43. Motzer RJ, Escudier B, Tomczak P, Hutson TE, Michaelson MD, Negrier S, *et al.* Axitinib versus sorafenib as second-line treatment for advanced renal cell carcinoma: overall survival analysis and updated results from a randomised phase 3 trial. *Lancet Oncol* 2013;**14**:552–62. https://doi.org/10.1016/S1470-2045(13)70093-7

44. Scottish Intermediate Guideline Network. *Scottish Intermediate Guideline Network (SIGN) Search Filters*. 2017. URL: www.sign.ac.uk/methodology/filters.html# (accessed 7 February 2017).

45. National Cancer Institute. *Cancer Therapy Evaluation Program: Common Terminology Criteria for Adverse Events (CTCAE)*. 2016. URL: https://ctep.cancer.gov/protocoldevelopment/electronic_applications/ctc.htm#ctc_40 (accessed 7 March 2017).

46. Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, *et al.* The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;**343**:d5928. https://doi.org/10.1136/bmj.d5928

47. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, *et al.* ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 2016;**355**:i4919. https://doi.org/10.1136/bmj.i4919

48. Dias S, Welton NJ, Sutton AJ, Ades AE. *NICE DSU Technical Support Document 2: A Generalised Linear Modelling Framework for Pairwise and Network Meta-analysis of Randomised Controlled Trials*. 2011 (updated 2016). URL: www.nice.dsu.org.uk/TSD2%20General%20meta%20analysis%20corrected%202Sep2016v2pdf (accessed 7 March 2017).

49. Dias S, Welton NJ, Sutton AJ, Ades AE. *NICE DSU Technical Support Document 1: Introduction to Evidence Synthesis for Decision Making*. 2011 (updated 2012). URL: www.nice.dsu.org.uk/TSD1%20Introductionfinal080512pdf (accessed 7 March 2017).

50. Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005;**331**:897–900. https://doi.org/10.1136/bmj.331.7521.897

51. Ades AE. A chain of evidence with mixed comparisons: models for multi-parameter synthesis and consistency of evidence. *Stat Med* 2003;**22**:2995–3016. https://doi.org/10.1002/sim.1566

52. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004;**23**:3105–24. https://doi.org/10.1002/sim.1875

53. Cella D, Michaelson MD, Bushmakin AG, Cappelleri JC, Charbonneau C, Kim ST, *et al.* Health-related quality of life in patients with metastatic renal cell carcinoma treated with sunitinib vs interferon-alpha in a phase III trial: final results and geographical analysis. *Br J Cancer* 2010;**102**:658–64. https://doi.org/10.1038/sj.bjc.6605552

54. Motzer RJ, Escudier B, McDermott DF, George S, Hammers HJ, Srinivas S, *et al.* Nivolumab versus Everolimus in Advanced Renal-Cell Carcinoma. *N Engl J Med* 2015;**373**:1803–13. https://doi.org/10.1056/NEJMoa1510665

55. Tannir NM, Jonasch E, Albiges L, Altinmakas E, Ng CS, Matin SF, *et al.* Everolimus versus sunitinib prospective evaluation in metastatic non-clear cell renal cell carcinoma (ESPN): a randomized multicenter phase 2 trial. *Eur Urol* 2016;**69**:866–74. https://doi.org/10.1016/j.eururo.2015.10.049

56. Eichelberg C, Vervenne WL, De Santis M, Fischer von Weikersthal L, Goebell PJ, Lerchenmüller C, *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 cancer. *Eur Urol* 2015;**68**:837–47. https://doi.org/10.1016/j.eururo.2015.04.017

57. Choueiri TK, Escudier B, Powles T, Tannir NM, Mainwaring PN, Rini BI, *et al.* Cabozantinib versus everolimus in advanced renal cell carcinoma (METEOR): final results from a randomised, open-label, phase 3 trial. *Lancet Oncol* 2016;**17**:917–27. https://doi.org/10.1016/S1470-2045(16)30107-3

58. Calvani N, Morelli F, Leo S, Orlando L, Lombardi L, Gnoni A, *et al.* Sequential use of sorafenib and sunitinib in advanced renal cell carcinoma: does the order of sequencing matter? *Med Oncol* 2012;**29**:1908–13. https://doi.org/10.1007/s12032-011-0048-0

59. Iacovelli R, Santini D, Rizzo M, Felici A, Santoni M, Verzoni E, *et al.* Bone metastases affect prognosis but not effectiveness of third-line targeted therapies in patients with metastatic renal cell carcinoma. *Can Urol Assoc J* 2015;**9**:263–7. https://doi.org/10.5489/cuaj.2377

60. Paglino C, Procopio G, Sabbatini R, Bellmunt J, Schmidinger M, Bearz A, *et al.* A retrospective analysis of two different sequences of therapy lines for advanced kidney cancer. *Anticancer Res* 2013;**33**:4999–5004.

61. Porta C, Procopio G, Carteni G, Sabbatini R, Bearz A, Chiappino I, *et al.* Sequential use of sorafenib and sunitinib in advanced renal-cell carcinoma (RCC): an Italian multicentre retrospective analysis of 189 patient cases. *BJU Int* 2011;**108**:E250–7. https://doi.org/10.1111/j.1464-410X.2011.10186.x

62. Vogelzang NJ, Pal SK, Signorovitch JE, Reichmann WM, Li N, Yang C, *et al.* Comparative effectiveness of everolimus and axitinib as second targeted therapies for metastatic renal cell carcinoma in the US: a retrospective chart review. *Curr Med Res Opin* 2016;**32**:741–7. https://doi.org/10.1185/03007995.2016.1140028

63. Wong MK, Yang H, Signorovitch JE, Wang X, Liu Z, Liu NS, *et al.* Comparative outcomes of everolimus, temsirolimus and sorafenib as second targeted therapies for metastatic renal cell carcinoma: a US medical record review. *Curr Med Res Opin* 2014;**30**:537–45. https://doi.org/10.1185/03007995.2013.871243

64. Motzer RJ, Escudier B, Oudard S, Hutson TE, Porta C, Bracarda S, *et al.* Efficacy of everolimus in advanced renal cell carcinoma: a double-blind, randomised, placebo-controlled phase III trial. *Lancet* 2008;**372**:449–56. https://doi.org/10.1016/S0140-6736(08)61039-9

65. Korhonen P, Zuber E, Branson M, Hollaender N, Yateman N, Katiskalahti T, *et al.* Correcting overall survival for the impact of crossover via a rank-preserving structural failure time (RPSFT) model in the RECORD-1 trial of everolimus in metastatic renal-cell carcinoma. *J Biopharm Stat* 2012;**22**:1258–71. https://doi.org/10.1080/10543406.2011.592233

66. Rini BI, Escudier B, Tomczak P, Kaprin A, Szczylik C, Hutson TE, *et al.* Comparative effectiveness of axitinib versus sorafenib in advanced renal cell carcinoma (AXIS): a randomised phase 3 trial. *Lancet* 2011;**378**:1931–9. https://doi.org/10.1016/S0140-6736(11)61613-9

67. Motzer RJ, Hutson TE, Cella D, Reeves J, Hawkins R, Guo J, *et al.* Pazopanib versus sunitinib in metastatic renal-cell carcinoma. *N Engl J Med* 2013;**369**:722–31. https://doi.org/10.1056/NEJMoa1303989

68. Cella D, Grunwald V, Nathan P, Doan J, Dastani H, Taylor F, *et al.* Quality of life in patients with advanced renal cell carcinoma given nivolumab versus everolimus in CheckMate 025: a randomised, open-label, phase 3 trial. *Lancet Oncol* 2016;**17**:994–1003. https://doi.org/10.1016/S1470-2045(16)30125-5

69. Cella D, Escudier B, Powles T, Tannir N, Donskov FN, Peltola K, *et al. Quality of life (QoL) in the Phase 3 METEOR Trial of Cabozantinib vs Everolimus for Advanced Renal Cell Carcinoma*. Poster presented at ESMO 18th World Congress on Gastrointestinal Cancer. Barcelona, 29 June–2 July 2016.

70. Beaumont JL, Butt Z, Baladi J, Motzer RJ, Haas T, Hollaender N, *et al.* Patient-reported outcomes in a phase iii study of everolimus versus placebo in patients with metastatic carcinoma of the kidney that has progressed on vascular endothelial growth factor receptor tyrosine kinase inhibitor therapy. *Oncologist* 2011;**16**:632–40. https://doi.org/10.1634/theoncologist.2010-0299

71. Beaumont J, Cella D, Hutson T. Patient-reported outcomes in a randomized trial of everolimus with metastatic renal cell carcinoma patients. *J Clin Oncol* 2009;**27**(Suppl. 15):e17516.

72. National Cancer Institute Division of Cancer Treatment and Diagnosis. *Common Terminology Criteria for Adverse Events v3.0 (CTCAE)*. 2006. URL: http://ctep.cancer.gov/protocolDevelopment/electronic_applications/docs/ctcaev3.pdf (accessed 23 May 2016).

73. National Cancer Institute Division of Cancer Treatment and Diagnosis. *Common Terminology Criteria for Adverse Events (CTCAE) Version 4.0*. Washington, DC: U.S. Department of Health and Human Services; 2010.

74. Calvo E, Escudier B, Motzer RJ, Oudard S, Hutson TE, Porta C, *et al.* Everolimus in metastatic renal cell carcinoma: subgroup analysis of patients with 1 or 2 previous vascular endothelial growth factor receptor-tyrosine kinase inhibitor therapies enrolled in the phase III RECORD-1 study. *Eur J Cancer* 2012;**48**:333–9. https://doi.org/10.1016/j.ejca.2011.11.027

75. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;**339**:b2535. https://doi.org/10.1136/bmj.b2535

76. Petrou P. Cost-effectiveness analysis of axitinib through a probabilistic decision model. *Expert Opin Pharmacother* 2015;**16**:1233–43. https://doi.org/10.1517/14656566.2015.1039982

77. Petrou PK, Talias MA. Cost-effectiveness of sorafenib compared to best supportive care in second line renal cell cancer from a payer perspective in Cyprus. *Expert Rev Pharmacoecon Outcomes Res* 2014;**14**:131–18. https://doi.org/10.1586/14737167.2014.873703

78. Petrou P, Talias MA. A pilot study to assess feasibility of value based pricing in Cyprus through pharmacoeconomic modelling and assessment of its operational framework: sorafenib for second line renal cell cancer. *Cost Eff Resour Alloc* 2014;**12**:12. https://doi.org/10.1186/1478-7547-12-12

79. Mihajlović J, Pechlivanoglou P, Sabo A, Tomić Z, Postma MJ. Cost-effectiveness of everolimus for second-line treatment of metastatic renal cell carcinoma in Serbia. *Clin Ther* 2013;**35**:1909–22. https://doi.org/10.1016/j.clinthera.2013.10.004

80. Casciano R, Chulikavit M, Di Lorenzo G, Liu Z, Baladi JF, Wang X, *et al.* Economic evaluation of everolimus versus sorafenib for the treatment of metastatic renal cell carcinoma after failure of first-line sunitinib. *Value Health* 2011;**14**:846–51. https://doi.org/10.1016/j.jval.2011.04.008

81. Hoyle M, Green C, Thompson Coon J, Liu Z, Welch K, Moxham T, *et al.* Cost-effectiveness of sorafenib for second-line treatment of advanced renal cell carcinoma. *Value Health* 2010;**13**:55–60. https://doi.org/10.1111/j.1524-4733.2009.00616.x

82. Thompson Coon J, Hoyle M, Green C, Liu Z, Welch K, Moxham T, *et al.* Bevacizumab, sorafenib tosylate, sunitinib and temsirolimus for renal cell carcinoma: a systematic review and economic evaluation. *Health Technol Assess* 2010;**14**(2). https://doi.org/10.3310/hta14020

83. Paz-Ares L, del Muro JG, Grande E, Diaz S. A cost-effectiveness analysis of sunitinib in patients with metastatic renal cell carcinoma intolerant to or experiencing disease progression on immunotherapy: perspective of the Spanish National Health System. *J Clin Pharm Ther* 2010;**35**:429–38. https://doi.org/10.1111/j.1365-2710.2009.01135.x

84. Purmonen T, Martikainen JA, Soini EJ, Kataja V, Vuorinen RL, Kellokumpu-Lehtinen PL. Economic evaluation of sunitinib malate in second-line treatment of metastatic renal cell carcinoma in Finland. *Clin Ther* 2008;**30**:382–92. https://doi.org/10.1016/j.clinthera.2008.02.013

85. Lopes M, Chulikavit M, Parikh R, Stern L, Liu Z, Rogerio J. Budget impact of everolimus in treating metastatic renal cell carcinoma. *Am J Pharm Benefits* 2012;**4**:SP41–8.

86. National Institute for Health and Care Excellence. *Guide to the Methods of Technology Appraisal 2013–5. The Reference Case 2013*. URL: www.nice.org.uk/process/pmg9/chapter/the-reference-case (accessed 4 April 2017).

87. Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, *et al.* Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technol Assess* 2004;**8**(36). https://doi.org/10.3310/hta8360

88. National Institute for Health and Care Excellence. *NICE Technology Appraisal Guidance TA219. Everolimus for the Second-line Treatment of Advanced Renal Cell Carcinoma*. London: NICE; 2009. URL: www.nice.org.uk/guidance/ta219 (accessed 1 December 2016).

89. National Institute for Health and Care Excellence. *NICE Technology Appraisal Guidance TA333. Axitinib for Treating Advanced Renal Cell Carcinoma After Failure of Prior Systemic Treatment*. London: NICE; 2015. URL: www.nice.org.uk/guidance/ta333 (accessed 1 December 2016).

90. National Institute for Health and Care Excellence. *NICE Technology Appraisal Guidance TA417. Nivolumab For Previously Treated Advanced Renal Cell Carcinoma*. London: NICE; 2016. URL: www.nice.org.uk/guidance/ta417 (accessed 6 December 2016).

91. Escudier B, Eisen T, Stadler WM, Szczylik C, Oudard S, Siebels M, *et al.* Sorafenib in advanced clear-cell renal-cell carcinoma. *N Engl J Med* 2007;**356**:125–34. https://doi.org/10.1056/NEJMoa060655

92. Motzer RJ, Michaelson MD, Redman BG, Hudes GR, Wilding G, Figlin RA, *et al.* Activity of SU11248, a multitargeted inhibitor of vascular endothelial growth factor receptor and platelet-derived growth factor receptor, in patients with metastatic renal cell carcinoma. *J Clin Oncol* 2006;**24**:16–24. https://doi.org/10.1200/JCO.2005.02.2574

93. Motzer RJ, Escudier B, Oudard S, Hutson TE, Porta C, Bracarda S, *et al.* Phase 3 trial of everolimus for metastatic renal cell carcinoma: final results and analysis of prognostic factors. *Cancer* 2010;**116**:4256–65. https://doi.org/10.1002/cncr.25219

94. Di Lorenzo G, Carteni G, Autorino R, Bruni G, Tudini M, Rizzo M, *et al.* Phase II study of sorafenib in patients with sunitinib-refractory metastatic renal cell cancer. *J Clin Onol* 2009;**27**:4469–74. https://doi.org/10.1200/JCO.2009.22.6480

95. Doyle S, Lloyd A, Walker M. Health state utility scores in advanced non-small cell lung cancer. *Lung Cancer* 2008;**62**:374–80. https://doi.org/10.1016/j.lungcan.2008.03.019

96. PenTAG. *Bevacizumab, Sorafenib Tosylate, Sunitinib, and Temsirolimus for Renal Cell Carcinoma: A Systematic Review and Economic Evaluation*. London: NICE; 2008.

97. Surveillance, Epidemiology and End Results. *Surveillance, Epidemiology and End Results (SEER) Program of the National Cancer Institute*. 2008 URL: www.seer.cancer.gov (accessed 23 June 2016).

98. US Department of Health & Human Services. *US Department of Health & Human Services. Medicare: The Official US Government Site for People with Medicare*. URL: www.medicare.gov/ (accessed 23 June 2016).

99. Motzer RJ, Michaelson MD, Rosenberg J, Bukowski RM, Curti BD, George DJ, *et al.* Sunitinib efficacy against advanced renal cell carcinoma. *J Urol* 2007;**178**:1883–7. https://doi.org/10.1016/j.juro.2007.07.030

100. Motzer RJ, Bacik J, Schwartz LH, Reuter V, Russo P, Marion S, *et al.* Prognostic factors for survival in previously treated patients with metastatic renal cell carcinoma. *J Clin Oncol* 2004;**22**:454–63. https://doi.org/10.1200/JCO.2004.06.132

101. Joint Formulary Committee. *British National Formulary*. 55 ed. London: BMJ Group and Pharmaceutical Press; 2008.

102. Joint Formulary Committee. *British National Formulary* (online). London: BMJ Group and Pharmaceutical Press. URL: www.medicinescomplete.com (September 2015).

103. Sherman S, Amzal B, Calvo E, Wang X, Park J, Liu Z, *et al.* An indirect comparison of everolimus versus axitinib in US patients with advanced renal cell carcinoma in whom prior sunitinib therapy failed. *Clin Ther* 2015;**37**:2552–9. https://doi.org/10.1016/j.clinthera.2015.09.013

104. Escudier B, Eisen T, Stadler WM, Szczylik C, Oudard S, Staehler M, *et al.* Sorafenib for treatment of renal cell carcinoma: final efficacy and safety results of the phase III treatment approaches in renal cancer global evaluation trial. *J Clin Oncol* 2009;**27**:3312–8. https://doi.org/10.1200/JCO.2008.19.5511

105. Wahlgren T, Harmenberg U, Sandström P, Lundstam S, Kowalski J, Jakobsson M, *et al.* Treatment and overall survival in renal cell carcinoma: a Swedish population-based study (2000-2008). *Br J Cancer* 2013;**108**:1541–9. https://doi.org/10.1038/bjc.2013.119

106. Sternberg CN, Davis ID, Mardiak J, Szczylik C, Lee E, Wagstaff 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. https://doi.org/10.1200/JCO.2009.23.9764

107. Escudier B, Roigas J, Gillessen S, Harmenberg U, Srinivas S, Mulder SF, *et al.* Phase II study of sunitinib administered in a continuous once-daily dosing regimen in patients with cytokine-refractory metastatic renal cell carcinoma. *J Clin Oncol* 2009;**27**:4068–75. https://doi.org/10.1200/JCO.2008.20.5476

108. Grünwald V, Weikert S, Pavel ME, Hörsch D, Lüftner D, Janni W, *et al.* Practical management of everolimus-related toxicities in patients with advanced solid tumors. *Onkologie* 2013;**36**:295–302. https://doi.org/10.1159/000350625

109. Sternberg C, Bracarda S, Carteni G, Lo Re G, Ruggeri EM, Masini C. Sunitinib expanded-access trial in metastatic renal cell carcinoma (mRCC) – final Italian results. *Eur J Cancer* 2013;**49**:S644.

110. Yang JC, Haworth L, Sherry RM, Hwu P, Schwartzentruber DJ, Topalian SL, *et al.* A randomized trial of bevacizumab, an anti-vascular endothelial growth factor antibody, for metastatic renal cancer. *N Engl J Med* 2003;**349**:427–34. https://doi.org/10.1056/NEJMoa021491

111. Hutson TE, Escudier B, Esteban E, Bjarnason GA, Lim HY, Pittman KB, *et al.* Randomized phase III trial of temsirolimus versus sorafenib as second-line therapy after sunitinib in patients with metastatic renal cell carcinoma. *J Clin Oncol* 2014;**32**:760–7. https://doi.org/10.1200/JCO.2013.50.3961

112. Motzer RJ, Rini BI, McDermott DF, Redman BG, Kuzel T, Harrison MR, *et al.* Nivolumab for metastatic renal cell carcinoma: results of a randomized phase II trial. *J Clin Oncol* 2015;**33**:1430–7. https://doi.org/10.1200/JCO.2014.59.0703

113. Motzer RJ, Nosov D, Eisen T, Bondarenko I, Lesovoy V, Lipatov O, *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. https://doi.org/10.1200/JCO.2012.47.4940

114. Royston P, Parmar MK. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. *Stats Med* 2002;**21**:2175–97. https://doi.org/10.1002/sim.1203

115. Addicott R, Dewar S. *Improving Choice at End of Life. A Descriptive Analysis of the Impact and Costs of the Marie Curie Delivering Choice Programme in Lincolnshire*. London: The King's Fund; 2008.

116. Guyot P, Ades AE, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Med Res Methodol* 2012;**12**:9. https://doi.org/10.1186/1471-2288-12-9

117. Jackson C. flexsurv: a platform for parametric survival modeling in R. *J Stat Software* 2016;**70**:1–33. https://doi.org/10.18637/jss.v070.i08

118. Latimer N. *NICE DSU Technical Support Document 14: Undertaking Survival Analysis for Economic Evaluations Alongside Clinical Trials – Extrapolation with Patient-level Data*. Sheffield: ScHARR, University of Sheffield; 2011.

119. Bushmakin A, Cappelleri JC, Korytowsky B, Sandin R, Matczak E, Cella D. *Sunitinib (SU) Dosing Schedule and Data Collection Timepoints: Impact on Quality of Life (QOL) Outcomes in Metastatic Renal Cell Carcinoma (MRCC)*. Poster presented at the 37th ESMO Congress, Vienna, 28 September–2 October 2012.

120. Cella D, Michaelson MD, Bushmakin AG, Cappelleri JC, Charbonneau C, Kim ST, *et al. Final Quality of Life (QOL) Results with Geographical Analysis for Sunitinib Versus Interferon-alfa as First-line Therapy in Patients with Metastatic Renal Cell Carcinoma (mRCC). European Journal of Cancer, Supplement*. Joint ECCO (European CanCer Organisation) 15–34th ESMO Multidisciplinary Congress, Berlin, 2009.

121. Escudier B, Rini BI, Hutson TE, Gore M, Oudard S, Tarazi J, *et al.* Updated results of the phase 3 AXIS trial: axitinib vs sorafenib as second-line therapy for metastatic renal cell carcinoma (mRCC). *Eur Urol* 2012;**11**:e81–ea. https://doi.org/10.1016/S1569-9056(12)60080-3

122. Cella D, Bushmakin AG, Cappelleri JC, Charbonneau C, Sandin R, Negrier S, *et al.* Comparison of health-related quality of life (HRQOL) in patients reporting the same adverse event (AE) fatigue on different treatments. *Eur J Cancer* 2011;**47**:S233. https://doi.org/10.1016/S0959-8049(11)71109-1

123. Tannir NM, Cohen L, Wang X, Thall P, Mathew PF, Jonasch E, *et al.* Improved tolerability and quality of life with maintained efficacy using twice-daily low-dose interferon-alpha-2b: results of a randomized phase II trial of low-dose versus intermediate-dose interferon-alpha-2b in patients with metastatic renal cell carcinoma. *Cancer* 2006;**107**:2254–61. https://doi.org/10.1002/cncr.22253

124. Rixe O, Bukowski RM, Michaelson MD, Wilding G, Hudes GR, Bolte O, *et al.* Axitinib treatment in patients with cytokine-refractory metastatic renal-cell cancer: a phase II study. *Lancet Oncol* 2007;**8**:975–84. https://doi.org/10.1016/S1470-2045(07)70285-1

125. Kröger MJ, Menzel T, Gschwend JE, Bergmann L. Life quality of patients with metastatic renal cell carcinoma and chemo-immunotherapy – a pilot study. *Anticancer Res* 1999;**19**:1553–5.

126. Elfiky A, Oh WK, Choueiri TK. Health-related quality of life and symptom improvement for patients with metastatic renal cell carcinoma: evaluating sunitinib and interferon alfa. *Am J Hematol* 2009;**8**.

127. Pickard AS, Cella D, Duh MS, Guerin A, Mishagina N, Antras L, *et al. Health-related Quality of Life in Patients with Advanced Renal Cell Carcinoma Receiving Pazopanib or Placebo in a Randomized Phase III Trial. J Clin Oncol*. ASCO Annual Meeting 2011, Chicago, IL, 2011. https://doi.org/10.1200/jco.2011.29.15_suppl.9096

128. De Groot S, Redekop W, Oosterwijk E, Kiemeney LC, Uyl-de Groot C. *Patient and Disease Characteristics are Important Determinants of Health-related Quality of Life of Patients with Metastatic Renal Cell Carcinoma Results from a Population-based Registry. Value in Health*. ISPOR (International Society for Pharmacoeconomics and Outcomes Research) 17th Annual European Congress, Amsterdam, the Netherlands, 8–12 November 2014.

129. Zbrozek AS, Hudes G, Levy D, Strahs A, Berkenblit A, DeMarinis R, Parasuraman S. Q-TWiST analysis of patients receiving temsirolimus or interferon alpha for treatment of advanced renal cell carcinoma. *PharmacoEconomics* 2010;**28**:577–84. https://doi.org/10.2165/11535290-000000000-00000

130. Cella D, Pickard AS, Duh MS, Guerin A, Mishagina N, Antràs L, *et al.* Health-related quality of life in patients with advanced renal cell carcinoma receiving pazopanib or placebo in a randomised phase III trial. *Eur J Cancer* 2012;**48**:311–23. https://doi.org/10.1016/j.ejca.2011.05.017

131. Karakiewicz PI, Nott L, Joshi A, Kannourakis G, Tarazi J, Alam M. Evaluation of response from axitinib per response evaluation criteria in solid tumors versus choi criteria in previously treated patients with metastatic renal cell carcinoma. *OncoTargets Therapy* 2016;**9**:2855–63.

132. Cella D, Beaumont JL. Pazopanib in the treatment of advanced renal cell carcinoma. *Ther Adv Urol* 2016;**8**:61–9. https://doi.org/10.1177/1756287215614236

133. Cella D, Escudier B, Rini B, Chen C, Bhattacharyya H, Tarazi J, *et al.* Patient-reported outcomes for axitinib vs sorafenib in metastatic renal cell carcinoma: phase III (AXIS) trial. *Br J Cancer* 2013;**108**:1571–8. https://doi.org/10.1038/bjc.2013.145

134. Kind P, Hardman G, Macran S. *UK Population Norms for EQ-5D*. York: University of York, Centre for Health Economics; 1999.

135. Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997;**35**:1095–108. https://doi.org/10.1097/00005650-199711000-00002

136. Department of Health. *NHS Reference Costs 2014–15*. London: Department of Health; 2015.

137. Edwards SJ, Osei-Assibey G, Berardi A, Karner C, Salih F, Bacelar M. Nivolumab for previously treated advanced or metastatic renal cell carcinoma: a single technology appraisal. London: *BMJ TAG*; 2015.

138. Curtis L, Burns A. *Unit Costs of Health and Social Care 2015*. Canterbury: Personal Social Services Research Unit, University of Kent; 2015.

139. Georghiou T, Bardsley M. *Exploring the Cost of Care at the End of Life*. London: The Nuffield Trust; 2014.

140. *Monthly Index of Medical Specialities*. 2016. URL: www.mims.co.uk (accessed 18 May 2017).

141. *Monthly Index of Medical Specialities*. www.mims.co.uk/drugs/respiratory-system/asthma-copd/flixotide-evohaler (accessed 6 December 2016).

142. Motzer RJ, Barrios CH, Kim TM, Falcon S, Cosgriff T, Harker WG, *et al.* Phase II randomized trial comparing sequential first-line everolimus and second-line sunitinib versus first-line sunitinib and second-line everolimus in patients with metastatic renal cell carcinoma. *J Clin Oncol* 2014;**32**:2765–72. https://doi.org/10.1200/JCO.2013.54.6911

143. NHS England. *Clinical Commissioning Policy: Palliative Radiotherapy for Bone Pain*. Leeds: NHS England; 2016.

144. Ouwens MJ, Philips Z, Jansen JP. Network meta-analysis of parametric survival curves. *Res Synth Method* 2010;**1**:258–71. https://doi.org/10.1002/jrsm.25

145. Choueiri TK, Escudier B, Powles T, Mainwaring PN, Rini BI, Donskov F, *et al.* Cabozantinib versus Everolimus in Advanced Renal-Cell Carcinoma. *N Engl J Med* 2015;**373**:1814–23. https://doi.org/10.1056/NEJMoa1510016

146. Motzer R, Escudier B, Tomczak P, Negrier S, Gore M, Tarazi J, *et al. Axitinib vs Sorafenib for Advanced Renal Cell Carcinoma: Phase III Overall Survival Results and Analysis of Prognostic Factors*. ESMO Congress 2012, Vienna, Austria, 2012.

147. Escudier B, Michaelson MD, Motzer RJ, Hutson TE, Clark JI, Lim HY, *et al.* Axitinib versus sorafenib in advanced renal cell carcinoma: subanalyses by prior therapy from a randomised phase III trial. *Br J Cancer* 2014;**110**:2821–8. https://doi.org/10.1038/bjc.2014.244

148. Rini BI, Quinn DI, Baum M, Wood LS, Tarazi J, Rosbrook B, *et al.* Hypertension among patients with renal cell carcinoma receiving axitinib or sorafenib: analysis from the randomized phase III AXIS trial. *Target Oncol* 2015;**10**:45–53. https://doi.org/10.1007/s11523-014-0307-z

149. Motzer R, Sharma P, McDermott D, George S, Hammers HJ, Srinivas S, *et al.* CheckMate 025 phase III trial: outcomes by key baseline factors and prior therapy for nivolumab (NIVO) versus everolimus (EVE) in advanced renal cell carcinoma (RCC). *J Clin Oncol*; 2016;**34**(Suppl. 2):498. https://doi.org/10.1200/jco.2016.34.2_suppl.498

150. Cella D, Gruenwald V, Nathan P, Doan J, Dastani H, Taylor F. Quality of life (QoL) and overall survival (OS) in patients (pts) with advanced clear-cell renal cell carcinoma (aRCC) treated with nivolumab (NIVI) vs. everolimus (EVE) in the phase III CheckMate 025 study. *J Clin Oncol* 2016;**34**(Supp. 15):4549.

151. Tannir NM, Jonasch E, Altinmakas E, Ng CS, Qiao W, Tamboli P, *et al.* Everolimus versus sunitinib prospective evaluation in metastatic non-clear cell renal cell carcinoma (The ESPN Trial): a multicenter randomized phase 2 trial. *J Clin Oncol* 2014;**32**(Suppl. 5):4505.

152. Porta C, Paglino C, Procopio G, Sabbatini R, Bellmunt J, Schmidinger M, *et al.* Optimizing the sequencial treatment of metastatic renal cell carcinoma (mRCC) – a retrospective, multicenter, analysis of 40 patients treated with either sorafenib, an mtor inhibitor (mTORI) and sunitinib, or sunitinib, an mTORI and sorafenib. *Eur J Cancer* 2011;**47**:S514. https://doi.org/10.1016/S0959-8049(11)72046-9

153. Porta C, Procopio G, Carteni G, Sabbatini R, Bearz A, Chiappino I, *et al.* An Italian multicentre retrospective analysis of 189 advanced renal-cell carcinoma patients treated sequentially, either with sorafenib followed by sunitinib, or vice versa. *Ann Oncol* 2010;**21**:viii289.

154. Escudier B, Ravaud A, Oudard S, Hutson T, Porta C, Bracarda S, *et al.* Phase-3 randomised trial of everolimus (RAD001) vs placebo in metastatic renal cell carcinoma. *Annals of Oncology* 2008;**19**(Suppl. 8):viii44–6.

155. Hutson T, Negrier S, Kay A. Randomized, placebo-controlled, phase 3 study of everolimus, a novel therapy for patients with metastatic renal cell carcinoma: subgroup analysis of patients progressing on prior bevacizumab therapy. *Eur J Cancer* 2009;**7**:434.

156. Wiederkehr D, Howe CJ, Casciano R, Motzer R, Zheng J, Baladi JF. 7131 Overall survival among metastatic renal cell carcinoma patients corrected for crossover using inverse probability of censoring weights: analyses from the everolimus phase III trial. *EJC Supplements* 2009;**7**:432. https://doi.org/10.1016/S1359-6349(09)71464-8

157. Kay A, Motzer RJ, Figlin RA, Escudier B, Oudard S, Porta C, *et al. Updated Data from a Phase 3 Trial of Everolimus (RAD001) Versus Placebo in Patients with Metastatic Renal Cell Carcinoma. European Urology, Supplements*. Conference: 24th Annual Congress of the European Association of Urology, EAU, Stockholm, Sweden, 2009.

158. White DA, Camus P, Endo M, Escudier B, Calvo E, Akaza H, *et al.* Noninfectious pneumonitis after everolimus therapy for advanced renal cell carcinoma. *Am J Resp Crit Care Med* 2010;**182**:396–403. https://doi.org/10.1164/rccm.200911-1720OC

159. Calvo E, Motzer RJ, Hutson TE, Oudard S, Porta C, Grunwald V, *et al.* Phase 3 record-1 study of everolimus in metastatic renal cell carcinoma (mRCC): subgroup analysis of patients (PTS) with 1 versus 2 prior vascular endothelial growth factor receptor-tyrosine kinase inhibitor (VEGFR-TKI) therapies. *Ann Oncol* 2010;**21**:viii285.

160. Osanto S, Hutson T, Calvo E, Escudier B, Oudard S, Porta C, *et al.* Efficacy and safety of everolimus in elderly patients (pts) with metastatic renal cell carcinoma (mRCC). *J Clin Oncol* 2010;**28**(Suppl. 1):4608. https://doi.org/10.1200/jco.2010.28.15_suppl.4608

161. Porta C, Calvo E, Climent MA, Vaishampayan U, Osanto S, Ravaud A, *et al.* Efficacy and safety of everolimus in elderly patients with metastatic renal cell carcinoma: an exploratory analysis of the outcomes of elderly patients in the RECORD-1 Trial. *Eur Urol* 2012;**61**:826–33. https://doi.org/10.1016/j.eururo.2011.12.057

162. Figlin RA, Motzer R, Hutson TE, Oudard S, Porta C, Grunwald V, *et al.* Subgroup analysis of the phase 3 RECORD-1 trial of everolimus in patients with metastatic renal cell carcinoma: 1 versus 2 prior vascular endothelial growth factor receptor-tyrosine kinase inhibitor therapies. *BJU Int* 2012;**109**:4–5.

163. Figlin RA, Calvo E, Motzer RJ, Hutson TE, Oudard S, Porta C, *et al.* Everolimus in metastatic renal cell carcinoma (mRCC): subgroup analysis of patients (pts) with one versus two prior vascular endothelial growth factor receptor tyrosine kinase inhibitor (VEGFR-TKI) therapies enrolled in the phase III RECORD-1 study. *J Clin Oncol* 2011;**29**(Suppl. 7):304. https://doi.org/10.1200/jco.2011.29.7_suppl.304

164. Oudard S, Thiam R, Fournier LS, Medioni J, Lamuraglia M, Scotte F, *et al.* Optimisation of the tumour response threshold in patients treated with everolimus for metastatic renal cell carcinoma: analysis of response and progression-free survival in the RECORD-1 study. *Eur J Cancer* 2012;**48**:1512–18. https://doi.org/10.1016/j.ejca.2012.01.027

165. Michel MS, Vervenne W, Goebell PJ, Von Weikersthal LF, Freier W, De Santis M, *et al.* Phase III randomized sequential open-label study to evaluate efficacy and safety of sorafenib (SO) followed by sunitinib (SU) versus sunitinib followed by sorafenib in patients with advanced/metastatic renal cell carcinoma without prior systemic therapy (SWITCH Study): safety interim analysis results. *J Clin Oncology* 2012;**30**(Suppl.15):4539.

166. Eichelberg C, Fischer Von Weikersthal L, Goebell P, Lerchenmuller C, Zimmermann U, Freier W, *et al.* Phase III randomized sequential open-label study to evaluate efficacy and safety of sorafenib (SO) followed by sunitinib (SU) vs. sunitinib followed by sorafenib in patients with advanced/metastatic renal cell carcinoma (mRCC) without prior systemic therapy (SWITCH Study) – safety interim analysis results. *Urologe – Ausgabe A*. 2012;**51**:35.

167. Pal SK, Jonasch E, Signorovitch JE, Reichmann WM, Li N, Liu Z, *et al.* Real-world dosing and drug costs with everolimus or axitinib as second targeted therapies for advanced renal cell carcinoma: a retrospective chart review in the US. *J Med Econ* 2016;**19**:462–8. https://doi.org/10.3111/13696998.2015.1131705

168. Kind P. The EuroQol Instrument: An Index of Health-Related Quality of Life. In Spiker B, editor. *Quality of Life and Pharmacoeconomics in Clinical Trials*. 2nd edn. Philadelphia, PA: Lippincott-Raven; 1996.

169. Cella D, Yount S, Du H, Dhanda R, Gondek K, Langefeld K, *et al.* Development and validation of the Functional Assessment of Cancer Therapy-Kidney Symptom Index (FKSI). *J Support Oncol* 2006;**4**:191–9.

170. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, *et al.* The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst* 1993;**85**:365–76. https://doi.org/10.1093/jnci/85.5.365

171. Motzer RJ, Rini BI, Bukowski RM, Curti BD, George DJ, Hudes GR, *et al.* Sunitinib in patients with metastatic renal cell carcinoma. *JAMA* 2006;**295**:2516–24. https://doi.org/10.1001/jama.295.21.2516

Appendix 1 Systematic searches of the literature

Clinical literature

Observational search

MEDLINE (via Ovid) – epub ahead of print, In-Process & Other Non-Indexed Citations, MEDLINE(R) (via Ovid) Daily and MEDLINE(R) (via Ovid) 1946 to present
Date range searched: inception to 22 June 2016.

Search strategy

#	Search terms	Results
1	Carcinoma, Renal Cell/	26,660
2	(renal cell carcinoma$ or cell renal carcinoma$ or renal carcinoma$ or kidney carcinoma$ or kidney cell carcinoma$ or renal adenocarcinoma$ or kidney adenocarcinoma$ or adenocarcinoma$renal or adenocarcinoma$kidney$).mp.	32,899
3	(hypernephroma$ or nephroid carcinoma$ or hypernephroid carcinoma$ or kidney hypernephroma$ or kidney pelvic carcinoma$ or kidney pyelocarcinoma$ or renal hypernephroma$ or grawitz tumo?r$ or renal cell neoplasm$ or renal cell cancer$ or renal tumo?r$ or carcinoma chromophobe cell kidney$ or chromophobe cell kidney carcinoma$).mp.	11,972
4	kidney neoplasms/	60,597
5	(cancer$ adj2 kidney$1).ti,ab.	3792
6	(neoplasm$1 adj2 kidney$1).ti,ab.	291
7	(neoplasm$1 adj2 renal).ti,ab	1539
8	(cancer$ adj2 renal).ti,ab.	8225
9	(tumo?r$1 adj2 kidney$1).ti,ab.	3481
10	(tumo?r$1 adj2 renal).ti,ab.	9801
11	or/1-10	76,421
12	(axitinib or inlyta or AG013736 or 'AG 013736').mp.	582
13	(sorafenib or nexavar or 'bay 43-9006' or 'bay 439006' or bay43-9006 or bay439006).mp.	5717
14	(sunitinib or sutent or pha 2909040ad or pha2909040ad or 'su 010398' or 'su 011248' or su 10398 or su10398 or su 11248 or su010398 or su011248 or su11248).mp.	4485
15	or/12-14	9023
16	Epidemiologic studies/	7168
17	exp case control studies/	794,256
18	exp cohort studies/	1,559,316
19	Case control.tw.	96,785
20	(cohort adj (study or studies)).tw.	122,494
21	Cohort analy$.tw.	5031
22	(Follow up adj (study or studies)).tw.	42,259
23	(observational adj (study or studies)).tw.	63,698
24	Longitudinal.tw.	177,454
25	Retrospective.tw.	355,615

#	Search terms	Results
26	Cross sectional.tw.	228,187
27	Cross-sectional studies/	219,662
28	or/16-27	2,266,798
29	11 and 15 and 28	521
30	case report.tw.	242,730
31	letter/	929,567
32	historical article/	333,071
33	or/30-32	1,492,135
34	29 not 33	505
35	Animals/ not Humans/	4,234,583
36	34 not 35	503
37	(editorial or letter).pt.	1,339,526
38	36 not 37	502

EMBASE (via Ovid)

Date range searched: inception to 22 June 2016.

Search strategy

#	Search terms	Results
1	Carcinoma, Renal Cell/	16,617
2	(renal cell carcinoma$ or cell renal carcinoma$ or renal carcinoma$ or kidney carcinoma$ or kidney cell carcinoma$ or renal adenocarcinoma$ or kidney adenocarcinoma$ or adenocarcinoma$renal or adenocarcinoma$kidney$).mp.	61,163
3	(hypernephroma$ or nephroid carcinoma$ or hypernephroid carcinoma$ or kidney hypernephroma$ or kidney pelvic carcinoma$ or kidney pyelocarcinoma$ or renal hypernephroma$ or grawitz tumo?r$ or renal cell neoplasm$ or renal cell cancer$ or renal tumo?r$ or carcinoma chromophobe cell kidney$ or chromophobe cell kidney carcinoma$).mp.	16,595
4	kidney neoplasms/	12,668
5	(cancer$ adj2 kidney$1).ti,ab.	5062
6	(neoplasm$1 adj2 kidney$1).ti,ab.	304
7	(neoplasm$1 adj2 renal).ti,ab.	2013
8	(cancer$ adj2 renal).ti,ab.	11,424
9	(tumo?r$1 adj2 kidney$1).ti,ab.	4511
10	(tumo?r$1 adj2 renal).ti,ab.	13,580
11	or/1-10	87,168
12	(axitinib or inlyta or AG013736 or 'AG 013736').mp.	2904
13	(sorafenib or nexavar or 'bay 43-9006' or 'bay 439006' or bay43-9006 or bay439006).mp.	20,539
14	(sunitinib or sutent or pha 2909040ad or pha2909040ad or 'su 010398' or 'su 011248' or su 10398 or su10398 or su 11248 or su010398 or su011248 or su11248).mp.	16,949
15	or/12-14	29,644
16	Clinical study/	122,871

#	Search terms	Results
17	Case control study	106,767
18	Family study/	11,456
19	Longitudinal study/	88,783
20	Retrospective study/	471,345
21	Prospective study/	338,477
22	Randomized controlled trials/	100,746
23	21 not 22	335,588
24	Cohort analysis/	246,874
25	(Cohort adj (study or studies)).mp.	168,118
26	(Case control adj (study or studies)).tw.	95,000
27	(follow up adj (study or studies)).tw.	52,541
28	(observational adj (study or studies)).tw.	92,382
29	(epidemiologic$ adj (study or studies)).tw.	86,434
30	(cross sectional adj (study or studies)).tw.	120,454
31	or/16-20,23-30	1,589,069
32	11 and 15 and 31	912
33	Animals/ not Humans/	1,164,794
34	32 not 33	912
35	(editorial or letter).pt.	1,454,340
36	34 not 35	894

Randomised controlled trial search

The Cochrane Library, CENTRAL, DARE, NHS EED

Date range searched: inception to 9 June 2016.

Search strategy

#	Search terms
1	MeSH descriptor: [Carcinoma, Renal Cell] explode all trees
2	('renal cell' next carcinoma*) or ('cell renal' next carcinoma*) or (renal next carcinoma*) or (kidney next carcinoma*) or ('kidney cell' next carcinoma*) or (renal next adenocarcinoma*) or (kidney next adenocarcinoma*) or (adenocarcinoma* next renal) or (adenocarcinoma* next kidney*)
3	(hypernephroma*) or (nephroid next carcinoma*) or (hypernephroid next carcinoma*) or (kidney next hypernephroma*) or ('kidney pelvic' next carcinoma*) or (kidney next pyelocarcinoma*) or (renal next hypernephroma*) or (grawitz next tumo*r*) or ('renal cell' next neoplasm*) or ('renal cell' next cancer*) or (renal next tumo*r*) or ('carcinoma chromophobe cell' next kidney*) or ('chromophobe cell kidney' next carcinoma*)
4	MeSH descriptor: [Kidney Neoplasms] explode all trees
5	kidney near/3 (cancer* or neoplasm*)
6	renal near/3 (cancer* or neoplasm*)
7	renal near/2 tumo*r*
8	kidney* near/2 tumo*r*

#	Search terms
9	#1 or #2 or #3 or #4 or #5 or #6 or #7 or #8
10	(axitinib or inlyta or AG013736 or 'AG 013736')
11	(sorafenib or nexavar or 'bay 43-9006' or 'bay 439006' or bay43-9006 or bay439006)
12	(sunitinib or sutent or 'pha 2909040ad' or pha2909040ad or 'su 010398' or 'su 011248' or 'su 10398' or su10398 or 'su 11248' or su010398 or su011248 or su11248)
13	(everolimus or afinitor or certican or zortress or nvp-rad-001 or rad-001 or 'rad 001a' or 'rad 001' or rad001 or rad001a or 'sdz rad')
14	(nivolumab or opdivo or ONO4538 or 'ONO 4538' or BMS936558 or 'BMS 936558' or MDX1106 or 'MDX 1106')
15	#10 or #11 or #12 or #13 or #14
16	#9 and #15

MEDLINE(R) In-Process & Other Non-Indexed Citations (via Ovid) and MEDLINE(R) (via Ovid) 1946 to present

Date range searched: inception to 13 January 2016.

Search strategy

#	Search terms	Results
1	Carcinoma, Renal Cell/	25,196
2	(renal cell carcinoma$ or cell renal carcinoma$ or renal carcinoma$ or kidney carcinoma$ or kidney cell carcinoma$ or renal adenocarcinoma$ or kidney adenocarcinoma$ or adenocarcinoma$renal or adenocarcinoma$kidney$).mp.	30,523
3	(hypernephroma$ or nephroid carcinoma$ or hypernephroid carcinoma$ or kidney hypernephroma$ or kidney pelvic carcinoma$ or kidney pyelocarcinoma$ or renal hypernephroma$ or grawitz tumo?r$ or renal cell neoplasm$ or renal cell cancer$ or renal tumo?r$ or carcinoma chromophobe cell kidney$ or chromophobe cell kidney carcinoma$).mp.	11,291
4	kidney neoplasms/	58,522
5	(cancer$ adj2 kidney$1).ti,ab.	3429
6	(neoplasm$1 adj2 kidney$1).ti,ab.	280
7	(neoplasm$1 adj2 renal).ti,ab.	1436
8	(cancer$ adj2 renal).ti,ab.	7619
9	(tumo?r$1 adj2 kidney$1).ti,ab.	3274
10	(tumo?r$1 adj2 renal).ti,ab.	9271
11	or/1-10	72,460
12	(axitinib or inlyta or AG013736 or 'AG 013736').mp.	492
13	(sorafenib or nexavar or bay 43-9006 or bay 439006 or bay43-9006 or bay439006).mp.	5086
14	(sunitinib or sutent or pha 2909040ad or pha2909040ad or 'su 010398' or 'su 011248' or su 10398 or su10398 or su 11248 or su010398 or su011248 or su11248).mp.	4050
15	(everolimus or afinitor or certican or zortress or nvp-rad-001 or rad-001 or rad 001a or rad001 or rad001a or sdz rad).mp.	4215
16	(nivolumab or opdivo or ONO4538 or ONO 4538 or BMS936558 or BMS 936558 or MDX1106 or MDX 1106).mp.	304
17	(temsirolimus or cci-779 or cell-cycle-inhibitor-779 or nsc 683864 or nsc683864 or torisel).mp.	1223
18	(bevacizumab or avastin or nsc 704865 or nsc704865 or anti-vegf or rhumab-vegf).mp.	13,848

#	Search terms	Results
19	(alpha-interferon or alfaferone or alferon or alpha ferone or cilferon or ginterferon or interferon-alpha or intropa or kemron or leukinferon or leukinferron or leukocyte interferon or refecon a or referon a3 or sumiferon or sumipheron or veldona).mp.	37,282
20	(armala or pazopanib or gw786034 or gw 786034 or sb 710468 or sb710468 or votrient).mp.	818
21	(biotest or bioleukin or interleukin-ii or 'interleukin-2 or il-2 or il2 or ro-236019 or tcgf or tsf).mp.	75,643
22	or/12-21	135,497
23	Randomized Controlled Trials as Topic/	99,847
24	randomized controlled trial/	403,636
25	Random Allocation/	84,835
26	Double Blind Method/	132,170
27	Single Blind Method/	21,076
28	clinical trial/	495,802
29	clinical trial, phase i.pt.	15,426
30	clinical trial, phase ii.pt.	24,957
31	clinical trial, phase iii.pt.	10,475
32	clinical trial, phase iv.pt.	1091
33	controlled clinical trial.pt.	89,944
34	randomized controlled trial.pt.	403,636
35	multicenter study.pt.	191,590
36	clinical trial.pt.	495,802
37	exp Clinical Trials as topic/	285,709
38	(clinical adj trial$).tw.	244,874
39	((singl$ or doubl$ or treb$ or tripl$) adj (blind$3 or mask$3)).tw.	138,299
40	PLACEBOS/	32,935
41	placebo$.tw.	170,679
42	randomly allocated.tw.	19,460
43	(allocated adj2 random$).tw.	22,196
44	or/23-43	1,265,584
45	case report.tw.	227,909
46	letter/	897,682
47	historical article/	325,111
48	or/45-47	1,438,121
49	44 not 48	1,234,879
50	11 and 22 and 49	2186
51	Animals/ not Humans/	4,137,434
52	50 not 51	2170
53	(editorial or letter).pt.	1,288,582
54	52 not 53	2148

EMBASE (via Ovid) 1974 to week 2 2016

Date range searched: inception to 13 January 2016.

Search strategy

#	Search terms	Results
1	Carcinoma, Renal Cell/	16,438
2	(renal cell carcinoma$ or cell renal carcinoma$ or renal carcinoma$ or kidney carcinoma$ or kidney cell carcinoma$ or renal adenocarcinoma$ or kidney adenocarcinoma$ or adenocarcinoma$renal or adenocarcinoma$kidney$).mp.	58,651
3	(hypernephroma$ or nephroid carcinoma$ or hypernephroid carcinoma$ or kidney hypernephroma$ or kidney pelvic carcinoma$ or kidney pyelocarcinoma$ or renal hypernephroma$ or grawitz tumo?r$ or renal cell neoplasm$ or renal cell cancer$ or renal tumo?r$ or carcinoma chromophobe cell kidney$ or chromophobe cell kidney carcinoma$).mp.	16,022
4	kidney neoplasms/	12,405
5	(cancer$ adj2 kidney$1).ti,ab.	4742
6	(neoplasm$1 adj2 kidney$1).ti,ab.	297
7	(neoplasm$1 adj2 renal).ti,ab.	1939
8	(cancer$ adj2 renal).ti,ab.	10,864
9	(tumo?r$1 adj2 kidney$1).ti,ab.	4339
10	(tumo?r$1 adj2 renal).ti,ab.	13,104
11	or/1-10	83,750
12	(axitinib or inlyta or AG013736 or 'AG 013736').mp.	2685
13	(sorafenib or nexavar or bay 43-9006 or bay 439006 or bay43-9006 or bay439006).mp.	19,312
14	(sunitinib or sutent or pha 2909040ad or pha2909040ad or 'su 010398' or 'su 011248' or su 10398 or su10398 or su 11248 or su010398 or su011248 or su11248).mp.	16,052
15	(everolimus or afinitor or certican or zortress or nvp-rad-001 or rad-001 or rad 001a or rad001 or rad001a or sdz rad).mp.	18,144
16	(nivolumab or opdivo or ONO4538 or ONO 4538 or BMS936558 or BMS 936558 or MDX1106 or MDX 1106).mp.	1401
17	(temsirolimus or cci-779 or cell-cycle-inhibitor-779 or nsc 683864 or nsc683864 or torisel).mp.	6443
18	(bevacizumab or avastin or nsc 704865 or nsc704865 or anti-vegf or rhumab-vegf).mp.	41,877
19	(alpha-interferon or alfaferone or alferon or alpha ferone or cilferon or ginterferon or interferon-alpha or introma or kemron or leukinferon or leukinferron or leukocyte interferon or refecon a or referon a3 or sumiferon or sumipheron or veldona).mp.	62,818
20	(armala or pazopanib or gw786034 or gw 786034 or sb 710468 or sb710468 or votrient).mp.	4196
21	(biotest or bioleukin or interleukin-ii or 'interleukin-2 or il-2 or il2 or ro-236019 or tcgf or tsf).mp.	118,163
22	or/12-21	247,337
23	Clinical trial/	855,321
24	Randomized controlled trial/	391,268
25	Randomization/	68,690
26	Single blind procedure/	21,252
27	Double blind procedure/	127,454
28	Crossover procedure/	45,414
29	Placebo/	280,430
30	Randomi?ed controlled trial$.tw.	127,639

#	Search terms	Results
31	Rct.tw.	19,055
32	Random allocation.tw.	1511
33	Randomly allocated.tw.	23,955
34	Allocated randomly.tw.	2092
35	(allocated adj2 random).tw.	827
36	Single blind$.tw.	16,967
37	Double blind$.tw.	163,789
38	((treble or triple) adj blind$).tw.	534
39	Placebo$.tw.	230,396
40	Prospective study/	316,440
41	or/23-40	1,544,641
42	Case study/	35,640
43	Case report.tw.	305,750
44	Abstract report/ or letter/	964,248
45	or/42-44	1,298,935
46	41 not 45	1,503,886
47	11 and 22 and 46	4590
48	Animals/ not Humans/	1,158,981
49	47 not 48	4587
50	(editorial or letter).pt.	1,414,836
51	49 not 50	4505

Economic literature search

Economic evaluation and costing studies

Ovid MEDLINE(R) 1946 to present
Date range searched: inception to 18 February 2016.

Search strategy

#	Search terms	Results
1	Carcinoma, Renal Cell/	25,300
2	(renal cell carcinoma$ or cell renal carcinoma$ or renal carcinoma$ or kidney carcinoma$ or kidney cell carcinoma$ or renal adenocarcinoma$ or kidney adenocarcinoma$ or adenocarcinoma$renal or adenocarcinoma$kidney$).tw.	30,947
3	hypernephroma$ or nephroid carcinoma$ or hypernephroid carcinoma$ or kidney hypernephroma$ or kidney pelvic carcinoma$ or kidney pyelocarcinoma$ or renal hypernephroma$ or grawitz tumo?r$ or renal cell neoplasm$ or renal cell cancer$ or renal tumo?r$ or carcinoma chromophobe cell kidney$ or chromophobe cell kidney carcinoma$).tw.	11,217
4	kidney neoplasms/	58,639
5	(cancer$ adj2 kidney$1).ti,ab.	3449

#	Search terms	Results
6	(neoplasm$1 adj2 kidney$1).ti,ab.	281
7	(neoplasm$1 adj2 renal).ti,ab.	1437
8	(cancer$ adj2 renal).ti,ab	7655
9	(tumo?r$1 adj2 kidney$1).ti,ab.	9304
10	(tumo?r$1 adj2 renal).ti,ab.	3286
11	or/1-10	72,592
12	(axitinib or ag013736 or inlyta).tw.	420
13	(tivozanib or av-951).tw.	55
14	(pazopanib or armala or gw786034 or sb710468).tw.	751
15	(alpha-interferon or alfaferone or alferon or alpha ferone or cilferon or ginterferon or interferon-alpha or introma or kemron or leukinferon or leukinferron or leukocyte interferon or refecon a or referon a3 or sumiferon or sumipheron or veldona).tw.	22,681
16	(biotest or bioleukin or interleukin-ii or interleukin-2 or il-2 or il2 or ro-236019 or tcgf or tsf).tw.	63,418
17	interleukin$.tw.	184,176
18	(sunitinib or sutent or pha 2909040ad or pha2909040ad or 'su 010398' or 'su 011248' or su 10398 or su10398 or su 11248 or su010398 or su011248 or su11248).tw.	3665
19	(sorafenib bay 43-9006 or bay 439006 or bay43-9006 or bay439006 or nexavar).tw.	180
20	(everolimus or afinitor or nvp-rad-001 or rad-001 or rad 001a or rad001 or rad001a or sdz rad).tw.	3816
21	(temsirolimus or cci-779 or cell-cycle-inhibitor-779 or nsc 683864 or nsc683864 or torisel).tw.	1075
22	(bevacizumab or avastin or nsc 704865 or nsc704865 or anti-vegf or rhumab-vegf).tw. (12737)	12,737
23	(nivolumab or opdivo or ONO4538 or ONO 4538 or BMS936558 or BMS 936558 or MDX1106 or MDX 1106).tw.	246
24	or/12-23	247,802
25	11 and 24	5581
26	Animals/ not Humans/	4,145,244
27	25 not 26	5390
28	economics/	26,626
29	exp 'costs and cost analysis'/	193,384
30	exp economics, hospital/	21,017
31	economics, medical/	8842
32	economics, pharmaceutical/	2600
33	(economic$ or pharmaeconomic$ or pharmacoeconomic$ or pharmaco-economic$).tw.	179,254
34	(cost or costs or costly or costing or costed).tw.	383,280
35	value for money.tw.	1126
36	(Quality-adjusted life year$ or QALY$).tw.	8546
37	or/28-36	632,157
38	27 and 37	**109**

EMBASE 1974 to present

Date range searched: inception to 18 February 2016.

Search strategy

#	Search terms	Results
1	kidney carcinoma/	51,430
2	(renal cell carcinoma$ or cell renal carcinoma$ or renal carcinoma$ or kidney carcinoma$ or kidney cell carcinoma$ or renal adenocarcinoma$ or kidney adenocarcinoma$ or adenocarcinoma$renal or adenocarcinoma$kidney$).tw.	42,721
3	(hypernephroma$ or nephroid carcinoma$ or hypernephroid carcinoma$ or kidney hypernephroma$ or kidney pelvic carcinoma$ or kidney pyelocarcinoma$ or renal hypernephroma$ or grawitz tumo?r$ or renal cell neoplasm$ or renal cell cancer$ or renal tumo?r$ or carcinoma chromophobe cell kidney$ or chromophobe cell kidney carcinoma$).tw.	15,457
4	kidney tumor/	31,122
5	(cancer$adj2 kidney$1).ti,ab.	4784
6	(neoplasm$1 adj2 kidney$1).ti,ab.	299
7	(neoplasm$1 adj2 renal).ti,ab.	1966
8	(cancer$adj2 renal).ti,ab.	10,949
9	(tumo?r$1 adj2 kidney$1).ti,ab.	4370
10	(tumo?r$1 adj2 renal).ti,ab.	13,184
11	or/1–10	91,731
12	axitinib/	2641
13	tivozanib/	349
14	pazopanib/	4149
15	alpha interferon/	47,530
16	interleukin 2/	76,944
17	sunitinib/	15,778
18	sorafenib/	18,955
19	everolimus/	17,546
20	temsirolimus/	6309
21	bevacizumab/	38,551
22	nivolumab/	1339
23	(axitinib or ag013736 or inlyta).tw.	1016
24	(tivozanib or av-951).tw.	208
25	(pazopanib or armala or gw786034 or sb710468).tw.	1622
26	(alpha-interferon or alfaferone or alferon or alpha ferone or cilferon or ginterferon or interferon-alpha or introma or kemron or leukinferon or leukinferron or leucocyte interferon or refecon a or referon a3 or sumiferon or sumipheron or veldona).tw.	27,507
27	(biotest or bioleukin or interleukin-ii or interleukin-2 or il-2 or il2 or ro-236019 or tcgf or tsf).tw.	78,247
28	interleukin$.tw.	216,386
29	(sunitinib or sutent or pha 2909040ad or pha2909040ad or 'su 010398' or 'su 011248' or su 10398 or su10398 or su 11248 or su010398 or su011248 or su11248).tw.	9651
30	(sorafenib bay 43–9006 or bay 439006 or bay43–9006 or bay439006 or nexavar).tw.	2825

#	Search terms	Results
31	(everolimus or afinitor or nvp-rad-001 or rad-001 or rad 001a or rad001 or rad001a or sdz rad).tw.	11,125
32	(temsirolimus or cci-779 or cell-cycle-inhibitor-779 or nsc 683864 or nsc683864 or torisel).tw.	3739
33	(bevacizumab or avastin or nsc 704865 or nsc704865 or anti-vegf or rhumab-vegf).tw.	26,280
34	(nivolumab or opdivo or ONO4538 or ONO 4538 or BMS936558 or BMS 936558 or MDX1106 or MDX 1106).tw.	859
35	or/12–34	399,561
36	11 and 35	15,443
37	Animals/not Humans/	1,158,981
38	36 not 37	15,421
39	Health Economics/	35,039
40	exp Economic Evaluation/	237,835
41	exp Health Care Cost/	228,726
42	pharmacoeconomics/	6240
43	(economic$or pharmaeconomic$or pharmacoeconomic$or pharmaco-economic$).tw.	234,407
44	(cost or costs or costly or costing or costed).tw.	510,756
45	value for money.tw.	1622
46	(Quality-adjusted life-year$or QALY$).tw.	14,234
47	quality-adjusted life-year/	15,511
48	or/39–47	890,255
49	38 and 48	616

The Cochrane Library

Date range searched: inception to 18 February 2016.

Search strategy

#	Search terms	Results
1	MeSH descriptor: [Carcinoma, Renal Cell] this term only	555
2	'renal cell carcinoma*' or 'cell renal carcinoma*' or 'renal carcinoma*' or 'kidney carcinoma*' or 'kidney cell carcinoma*' or 'renal adenocarcinoma*' or 'kidney adenocarcinoma*' or 'adenocarcinoma*renal' or 'adenocarcinoma*kidney*':ti,ab,kw (Word variations have been searched)	1026
3	hypernephroma* or 'nephroid carcinoma*' or 'hypernephroid carcinoma*' or 'kidney hypernephroma*' or 'kidney pelvic carcinoma*' or 'kidney pyelocarcinoma*' or 'renal hypernephroma*' or 'grawitz tumour*' or 'grawitz tumour*' or 'renal cell neoplasm*' or 'renal cell cancer*' or 'renal tumour*' or 'renal tumour*' or 'carcinoma chromophobe cell kidney*' or 'chromophobe cell kidney carcinoma*':ti,ab,kw (Word variations have been searched)	226
4	MeSH descriptor: [Kidney Neoplasms] this term only	715
5	cancer* near/2 kidney*:ti,ab,kw (Word variations have been searched)	188
6	neoplasm* near/2 kidney*:ti,ab,kw (Word variations have been searched)	783
7	neoplasm* near/2 renal:ti,ab,kw (Word variations have been searched)	14
8	cancer* near/2 renal:ti,ab,kw (Word variations have been searched)	309
9	(tumour* or tumour*) near/2 kidney*:ti,ab,kw (Word variations have been searched)	68

#	Search terms	Results
10	(tumour* or tumour*) near/2 renal:ti,ab,kw (Word variations have been searched)	82
11	#1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10	1620
12	axitinib or ag013736 or inlyta:ti,ab,kw (Word variations have been searched)	83
13	tivozanib or av-951:ti,ab,kw (Word variations have been searched)	25
14	pazopanib or armala or gw786034 or sb710468:ti,ab,kw (Word variations have been searched)	135
15	alpha-interferon or alfaferone or alferon or 'alpha ferone' or cilferon or ginterferon or interferon-alpha or introma or kemron or leukinferon or leukinferron or 'leucocyte interferon' or 'refecon a' or 'referon a3' or sumiferon or sumipheron or veldona:ti,ab,kw (Word variations have been searched)	4641
16	biotest or bioleukin or interleukin-ii or interleukin-2 or il-2 or il2 or ro-236019 or tcgf or tsf:ti,ab,kw (Word variations have been searched)	2879
17	interleukin*:ti,ab,kw (Word variations have been searched)	11,735
18	sunitinib or sutent or 'pha 2909040ad' or pha2909040ad or 'su 010398' or 'su 011248' or 'su 10398' or su10398 or 'su 11248' or su010398 or su011248 or su11248:ti,ab,kw (Word variations have been searched)	375
19	'sorafenib bay 43–9006' or 'bay 439006' or bay43–9006 or bay439006 or nexavar:ti,ab,kw (Word variations have been searched)	20
20	everolimus or afinitor or nvp-rad-001 or rad-001 or 'rad 001a' or rad001 or rad001a or 'sdz rad':ti,ab,kw (Word variations have been searched)	1341
21	temsirolimus or cci-779 or cell-cycle-inhibitor-779 or 'nsc 683864' or nsc683864 or torisel:ti,ab,kw (Word variations have been searched)	122
22	bevacizumab or avastin or 'nsc 704865' or nsc704865 or anti-vegf or rhumab-vegf:ti,ab,kw (Word variations have been searched)	2030
23	nivolumab or opdivo or ONO4538 or 'ONO 4538' or BMS936558 or 'BMS 936558' or MDX1106 or 'MDX 1106':ti,ab,kw (Word variations have been searched)	38
24	#12 or #13 or #14 or #15 or #16 or #17 or #18 or #19 or #20 or #21 or #22 or #23	20,118
25	#11 and #24	715
26	#25 and NHS EED	12
28	#25 and DARE	13

Quality of life

MEDLINE

Date range searched: inception to 18 July 2016.

Search strategy

#	Search terms	Results
1	Carcinoma, Renal Cell/	26,753
2	(renal cell carcinoma$or cell renal carcinoma$or renal carcinoma$or kidney carcinoma$or kidney cell carcinoma$or renal adenocarcinoma$or kidney adenocarcinoma$or adenocarcinoma$renal or adenocarcinoma$kidney$).mp.	29,403
3	(hypernephroma$or nephroid carcinoma$or hypernephroid carcinoma$or kidney hypernephroma$or kidney pelvic carcinoma$or kidney pyelocarcinoma$or renal hypernephroma$or grawitz tumo?r$or renal cell neoplasm$or renal cell cancer$or renal tumo?r$or carcinoma chromophobe cell kidney$or chromophobe cell kidney carcinoma$).mp.	10,882
4	kidney neoplasms/	60,749

#	Search terms	Results
5	(cancer$adj2 kidney$1).ti,ab.	3306
6	(neoplasm$1 adj2 kidney$1).ti,ab.	270
7	(neoplasm$1 adj2 renal).ti,ab.	1368
8	(cancer$adj2 renal).ti,ab.	7384
9	(tumo?r$1 adj2 kidney$1).ti,ab.	3188
10	(tumo?r$1 adj2 renal).ti,ab.	8936
11	or/1–10	71,406
12	exp quality of life/	140,401
13	(life adj2 qualit$3).ti,ab.	166,784
14	((quality adj2 life) or life quality or QOL or QoL).ti,ab.	168,184
15	(HQL or HRQL or HRQOL or HRQol).ti,ab.	11,380
16	(value adj2 life).ti,ab. or Value of Life/	5989
17	(quality-adjusted life-year$1 or QALY$or qaly$or quality-adjusted life-year$1).ti,ab. or Quality-Adjusted Life-years/	11,933
18	(daly$or DALY$).ti,ab.	1515
19	(disabilit$3 adj2 life).ti,ab.	2472
20	health status indicators/	21,626
21	(sf36 or sf-36 or sf 36 or short form 36 or shortform 36 or sf thirtysix or sf thirty six or shortform thirtysix or shortform thirty six or short form thirty six or short form thirtysix or short form thirty six).tw.	17,707
22	(sf6 or sf 6 or sf-6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six).tw.	1086
23	(sf6d or sf 6d or sf-6d or short form 6d or shortform 6d or sf six dimension$1 or short form six dimension$1).tw	511
24	(sf12 or sf 12 or sf-12 or short form 12 or shortform 12 or sf twelve of sftwelve or shortform twelve or short form twelve).tw.	3354
25	(sf16 or sf 16 or sf-16 or short form 16 or shortform 16 or sf sixteen or sfsixteen or shortform sixteen or short form sixteen).tw.	22
26	(sf20 or sf 20 or sf-20 or short form 20 or shortform 20 or sf twenty of sftwenty or shortform twenty of short form twenty).tw.	340
27	(euroqol or euro qol or eq5d or eq 5d or eq-5d or EQ5D or EQ-5D).tw.	5077
28	(hye or hyes or health$year$equivalent$).tw.	62
29	hui.tw.	755
30	(EORTC* or eortc* or European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30).tw.	80
31	(QLQ* or Quality of Life Questionnaire).tw.	7249
32	(standard gamble$or SG).tw.	6539
33	(time trade off or time tradeoff or TTO or time trade-off).tw.	1283
34	discrete choice experiment$.ti,ab.	574
35	(visual analogue$3 scale or VAS).tw.	42,701
36	((health stat$2 utilit$) or (health stat$2 value$) or (health stat$2 preference$) or HSUV).tw.	558
37	(health adj3 (utilit$3 or value$2 or preference$2)).tw.	7737
38	(person$trade-off or person$trade off or PTO).ti,ab.	558

#	Search terms	Results
39	(Contingent value or contingent valuation).ti,ab.	463
40	((quality adj3 wellbeing index) or (quality adj3 well-being index) or QWB).ti,ab.	195
41	(health utilit$index or HUI).ti,ab.	1196
42	disutilit$.tw.	259
43	((quality of well-being) or (quality of wellbeing) or (quality of well-being)).tw.	356
44	or/12–43	298,364
45	letter.pt.	893,981
46	editorial.pt.	384,816
47	comment.pt.	635,315
48	or/45–47	1,415,173
49	44 not 48	287,833
50	49 and 11	767

EMBASE 1974 to present

Date range searched: inception to 18 July 2016.

Search strategy

#	Search terms	Results
1	Carcinoma, Renal Cell/	16,629
2	(renal cell carcinoma$or cell renal carcinoma$or renal carcinoma$or kidney carcinoma$or kidney cell carcinoma$or renal adenocarcinoma$or kidney adenocarcinoma$or adenocarcinoma$renal or adenocarcinoma$kidney$).tw.	61,331
3	(hypernephroma$or nephroid carcinoma$or hypernephroid carcinoma$or kidney hypernephroma$or kidney pelvic carcinoma$or kidney pyelocarcinoma$or renal hypernephroma$or grawitz tumo?r$or renal cell neoplasm$or renal cell cancer$or renal tumo?r$or carcinoma chromophobe cell kidney$or chromophobe cell kidney carcinoma$).tw.	16,634
4	kidney neoplasms/	12,681
5	(cancer$adj2 kidney$1).ti,ab.	5084
6	(neoplasm$1 adj2 kidney$1).ti,ab.	304
7	(neoplasm$1 adj2 renal).ti,ab.	2017
8	(cancer$adj2 renal).ti,ab.	11,460
9	(tumo?r$1 adj2 kidney$1).ti,ab.	4533
10	(tumo?r$1 adj2 renal).ti,ab.	13,609
11	exp quality of life/	351,935
12	(life adj2 qualit$3).ti,ab.	274,643
13	((quality adj2 life) or life quality or QOL or QoL).ti,ab.	279,409
14	(HQL or HRQL or HRQOL or HRQol).ti,ab.	18,919
15	(value adj2 life).ti,ab. or Value of Life/	209,293
16	(quality-adjusted life-year$1 or QALY$or qaly$or quality-adjusted life-year$1).ti,ab. or Quality-Adjusted Life-years/	12,891
17	daly.ti,ab.	1365

#	Search terms	Results
18	(disabilit$3 adj2 life).ti,ab.	3581
19	exp health status indicators/	18,051
20	(sf36 or sf-36 or sf 36 or short form 36 or shortform 36 or sf thirtysix or sf thirty six or shortform thirtysix or shortform thirty six or short form thirty six or short form thirtysix or short form thirty six).tw.	29,509
21	(sf6 or sf 6 or sf-6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six).tw.	1749
22	(sf6d or sf 6d or sf-6d or short form 6d or shortform 6d or sf six dimension$1 or short form six dimension$1).tw	991
23	(sf12 or sf 12 or sf-12 or short form 12 or shortform 12 or sf twelve of sftwelve or shortform twelve or short form twelve).tw.	5983
24	(sf16 or sf 16 or sf-16 or short form 16 or shortform 16 or sf sixteen or sfsixteen or shortform sixteen or short form sixteen).tw.	41
25	(sf20 or sf 20 or sf-20 or short form 20 or shortform 20 or sf twenty of sftwenty or shortform twenty of short form twenty).tw.	369
26	(euroqol or euro qol or eq5d or eq 5d or eq-5d or EQ5D or EQ-5D).tw.	10,102
27	(hye or hyes or health$year$equivalent$).tw.	116
28	hui.tw.	1173
29	(EORTC or eortc or European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30).tw.	122
30	(QLQ or Quality of Life Questionnaire).tw.	12,749
31	(standard gamble$or SG).tw.	10,005
32	(time trade off or time tradeoff or TTO or time trade-off).tw.	1874
33	discrete choice experiment$.ti,ab.	1004
34	(visual analogue$3 scale or VAS).tw.	72,785
35	((health stat$2 utilit$) or (health stat$2 value$) or (health stat$2 preference$) or HSUV).tw.	1015
36	(health adj3 (utilit$3 or value$2 or preference$2)).tw.	11,478
37	(person$trade-off or person$trade off or PTO).ti,ab.	723
38	(Contingent value or contingent valuation).ti,ab.	670
39	((quality adj3 wellbeing index) or QWB).ti,ab.	215
40	(health utilit$index or HUI).ti,ab.	1745
41	disutilit$.tw.	506
42	((quality of well-being) or (quality of wellbeing) or (quality of well-being)).tw.	420
43	letter.pt.	904,099
44	editorial.pt.	486,431
45	comment.pt.	0
46	or/1–10	85,097
47	or/43–45	1,390,530
48	46 not 47	82,455
49	or/11–42	586,754
50	48 and 49	1805

CENTRAL

Date range searched: inception to 18 July 2016.

Search strategy

#	Search terms	Results
1	MeSH descriptor: [Carcinoma, Renal Cell] this term only in Trials	546
2	'renal cell carcinoma*' or 'cell renal carcinoma*' or 'renal carcinoma*' or 'kidney carcinoma*' or 'kidney cell carcinoma*' or 'renal adenocarcinoma*' or 'kidney adenocarcinoma*' or 'adenocarcinoma*renal' or 'adenocarcinoma*kidney*':ti,ab,kw (Word variations have been searched) in Trials	1093
3	hypernephroma* or 'nephroid carcinoma*' or 'hypernephroid carcinoma*' or 'kidney hypernephroma*' or 'kidney pelvic carcinoma*' or 'kidney pyelocarcinoma*' or 'renal hypernephroma*' or 'grawitz tumour*' or 'grawitz tumour*' or 'renal cell neoplasm*' or 'renal cell cancer*' or 'renal tumour*' or 'renal tumour*' or 'carcinoma chromophobe cell kidney*' or 'chromophobe cell kidney carcinoma*':ti,ab,kw (Word variations have been searched) in Trials	234
4	MeSH descriptor: [Kidney Neoplasms] this term only in Trials	720
5	cancer* near/2 kidney*:ti,ab,kw (Word variations have been searched) in Trials	200
6	neoplasm* near/2 kidney*:ti,ab,kw (Word variations have been searched) in Trials	795
7	neoplasm* near/2 renal:ti,ab,kw (Word variations have been searched) in Trials	14
8	cancer* near/2 renal:ti,ab,kw (Word variations have been searched) in Trials	323
9	(tumour* or tumour*) near/2 kidney*:ti,ab,kw (Word variations have been searched) in Trials	72
10	(tumour* or tumour*) near/2 renal:ti,ab,kw (Word variations have been searched) in Trials	87
11	#1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10	1706
12	MeSH descriptor: [Quality of Life] explode all trees	17,827
13	MeSH descriptor: [Quality-Adjusted Life-years] explode all trees	4088
14	quality near/3 life:ti,ab,kw in Trials	46,742
15	qol:ti,ab,kw in Trials	6183
16	hrqol or hr qol or hrql or hr ql:ti,ab,kw in Trials	2553
17	QALY or quality-adjusted life-year or quality-adjusted life-year:ti,ab,kw in Trials	5335
18	SF 6d or SF-6d or sf6d or short form 6d or short form six dimension*:ti,ab,kw in Trials	149
19	SF 36 or SF-36 or SF36 or short form 36 or short form thirty six:ti,ab,kw in Trials	5318
20	eq-5d or eq5d or eq 5d or euroqol:ti,ab,kw in Trials	1983
21	hui or health utilities index:ti,ab,kw in Trials	185
22	standard gamble:ti,ab,kw in Trials	89
23	time trade off or TTO or time trade-off:ti,ab,kw in Trials	143
24	utilit*:ti,ab,kw in Trials	7803
25	#12 or #13 or #14 or #15 or #16 or #17 or #18 or #19 or #20 or #21 or #22 or #23 or #24in Trials	54,827
26	#11 and #25 in Trials	159

NHS EED and HTA

Date range searched: inception to 18 July 2016.

Search strategy

#	Search terms	Results
1	MeSH descriptor: [Carcinoma, Renal Cell] this term only	546
2	'renal cell carcinoma*' or 'cell renal carcinoma*' or 'renal carcinoma*' or 'kidney carcinoma*' or 'kidney cell carcinoma*' or 'renal adenocarcinoma*' or 'kidney adenocarcinoma*' or 'adenocarcinoma*renal' or 'adenocarcinoma*kidney*':ti,ab,kw (Word variations have been searched)	1093
3	hypernephroma* or 'nephroid carcinoma*' or 'hypernephroid carcinoma*' or 'kidney hypernephroma*' or 'kidney pelvic carcinoma*' or 'kidney pyelocarcinoma*' or 'renal hypernephroma*' or 'grawitz tumour*' or 'grawitz tumour*' or 'renal cell neoplasm*' or 'renal cell cancer*' or 'renal tumour*' or 'renal tumour*' or 'carcinoma chromophobe cell kidney*' or 'chromophobe cell kidney carcinoma*':ti,ab,kw (Word variations have been searched)	234
4	MeSH descriptor: [Kidney Neoplasms] this term only	720
5	cancer* near/2 kidney*:ti,ab,kw (Word variations have been searched)	200
6	neoplasm* near/2 kidney*:ti,ab,kw (Word variations have been searched)	795
7	neoplasm* near/2 renal:ti,ab,kw (Word variations have been searched)	14
8	cancer* near/2 renal:ti,ab,kw (Word variations have been searched)	323
9	(tumour* or tumour*) near/2 kidney*:ti,ab,kw (Word variations have been searched)	72
10	(tumour* or tumour*) near/2 renal:ti,ab,kw (Word variations have been searched)	87
11	#1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10	1706
12	#11 and Technology Assessments	74
13	#11 and Economic evaluations	42

Appendix 2 Inclusion and exclusion criteria for economic systematic reviews

Economic evaluation and costing studies

Inclusion and exclusion criteria applied in the economic evaluation systematic review.

Inclusion criteria	Exclusion criteria
• Full economic evaluations (cost-effectiveness, cost–utility, cost–benefit, cost-consequence or cost minimisation) • Costing/resource studies in RCC (for resource use review) • Any setting (to be as inclusive as possible) • At least one of the interventions or comparators as per the protocol (axitinib, everolimus, nivolumab, sorafenib, sunitinib)	• Abstracts with insufficient methodological details • Systematic reviews • Studies not available in the English language

Quality of life

Inclusion and exclusion criteria for the HRQoL systematic review.

Inclusion criteria	Exclusion criteria
Q1: possible generic, preference-based measure of HRQoL (e.g. EQ-5D, SF-6D, HUI) or standard gamble/time trade-off studies any setting (to be as inclusive as possible) Q2: possible generic, non-preference-based measure of HRQoL (e.g. SF-36) Q3: possible condition-specific measure of HRQoL	Abstracts with insufficient methodological details, systematic review, secondary source of utility value[a]

EQ-5D, European Quality of Life – 5 Dimensions; HUI, health utilities index; SF-36, Short Form questionnaire-36 items; SF-6D, Short Form questionnaire-6 Dimensions.

a Relevant systematic reviews and secondary sources of utility values were used as a source of additional studies for consideration.

Appendix 3 List of all included studies and the associated publications

Study	Reference
AXIS	Rini *et al.*[66]
	Motzer *et al.*[146]
	Motzer *et al.*[43]
	Cella *et al.*[133]
	Escudier *et al.*[147]
	Rini *et al.*[148]
Calvani *et al.*, 2013	Calvani *et al.*[58]
CheckMate 025	Motzer *et al.*[54]
	Motzer *et al.*[149]
	Cella *et al.*[150]
	Cella *et al.*[68]
ESPN	Tannir *et al.*[151]
	Tannir *et al.*[55]
Iacovelli *et al.*, 2015	Iacovelli *et al.*[59]
METEOR	Choueiri *et al.*[145]
	Choueiri *et al.*[57]
	Cella *et al.*[69]
	NICE[28]
Paglino *et al.*, 2013	Paglino *et al.*[60]
	Porta *et al.*[152]
Porta *et al.*, 2011	Porta *et al.*[153]
	Porta *et al.*[61]
RECORD-1	Motzer *et al.*[64]
	Escudier *et al.*[154]
	Hutson *et al.*[155]
	Wiederkehr *et al.*[156]
	Kay *et al.*[157]
	Motzer *et al.*[93]
	White *et al.*[158]
	Osanto *et al.*[160]
	Calvo *et al.*[159]
	Beaumont *et al.*[70]
	Calvo *et al.*[74]
	Korhonen *et al.*[65]
	Porta *et al.*[161]
	Figlin 2012[162]
	Figlin *et al.*[163]
	Oudard *et al.*[164]

Study	Reference
SWITCH	Michel *et al.*[165]
	Eichelberg *et al.*[166]
	Eichelberg *et al.*[56]
Vogelzang *et al.*, 2016	Vogelzang *et al.*[62]
	Pal *et al.*[167]
Wong *et al.*, 2014	Wong *et al.*[63]

Appendix 4 Results of the overall survival and progression-free survival sensitivity analyses

Progression-free survival sensitivity analysis 1

Hazard ratio and associated credible interval; hazard ratio < 1 favours treatment in left-hand column

	Treatment				
Treatment	BSC	Sunitinib	Axitinib	Cabozantinib	Everolimus
Everolimus	0.33 (0.25 to 0.43)	1.25 (0.91 to 1.75)	1.07 (0.85 to 1.37)	1.95 (1.59 to 2.42)	–
Cabozantinib	0.17 (0.12 to 0.24)	0.63 (0.44 to 0.95)	0.54 (0.40 to 0.76)	–	0.51 (0.41 to 0.63)
Axitinib	0.31 (0.21 to 0.44)	1.16 (0.85 to 1.63)	–	1.85 (1.32 to 2.51)	0.94 (0.73 to 1.18)
Sunitinib	0.27 (0.17 to 0.40)	–	0.86 (0.61 to 1.18)	1.58 (1.06 to 2.27)	0.80 (0.57 to 1.10)
BSC	–	3.75 (2.49 to 5.88)	3.21 (2.29 to 4.70)	6.05 (4.23 to 8.40)	3.06 (2.32 to 3.96)

Cells highlighted in green indicate statistically significant results.

Progression-free survival sensitivity analysis 2

Hazard ratio and associated credible interval; hazard ratio < 1 favours treatment in left-hand column

	Treatment				
Treatment	BSC	Sunitinib	Axitinib	Cabozantinib	Everolimus
Everolimus	0.33 (0.25 to 0.43)	1.28 (0.97 to 1.72)	1.07 (0.84 to 1.37)	1.95 (1.59 to 2.42)	–
Cabozantinib	0.17 (0.12 to 0.24)	0.65 (0.47 to 0.94)	0.54 (0.40 to 0.75)	–	0.51 (0.41 to 0.63)
Axitinib	0.31 (0.21 to 0.44)	1.20 (0.91 to 1.60)	––	1.85 (1.33 to 2.51)	0.94 (0.73 to 1.19)
Sunitinib	0.26 (0.17 to 0.38)	–	0.84 (0.62 to 1.10)	1.54 (1.06 to 2.15)	0.78 (0.58 to 1.03)
BSC	–	3.85 (2.63 to 5.84)	3.20 (2.29 to 4.67)	6.05 (4.23 to 8.40)	3.06 (2.32 to 3.96)

Cells highlighted in green indicate statistically significant results.

Overall survival sensitivity analysis 1

Hazard ratio and associated credible interval; hazard ratio < 1 favours treatment in left-hand column

	Treatment				
Treatment	BSC	Axitinib	Nivolumab	Cabozantinib	Everolimus
Everolimus	0.53 (0.22 to 1.64)	0.74 (0.56 to 0.99)	1.36 (1.12 to 1.67)	1.51 (1.22 to 1.90)	–
Cabozantinib	0.34 (0.14 to 1.11)	0.48 (0.34 to 0.71)	0.89 (0.67 to 1.22)	–	0.66 (0.53 to 0.82)
Nivolumab	0.38 (0.16 to 1.22)	0.54 (0.38 to 0.77)	–	1.12 (0.82 to 1.49)	0.73 (0.60 to 0.89)
Axitinib	0.93 (0.29 to 2.31)	–	1.87 (1.29 to 2.60)	2.07 (1.41 to 2.92)	1.36 (1.01 to 1.78)
BSC	–	1.08 (0.43 to 3.50)	2.62 (0.82 to 6.35)	2.90 (0.90 to 7.10)	1.90 (0.61 to 4.55)

Cells highlighted in green indicate statistically significant results.

Appendix 5 Mixed-treatment comparison model characteristics and inconsistency assessments

Summary of mixed-treatment comparison model characteristics for progression-free survival analyses

	PFS		
Characteristic	Primary analysis	SA1	SA2
Total residual deviance	2	8	8
Number of data points	2	8	10

Results of the assessments for inconsistency in the data loops in mixed-treatment comparison sensitivity analysis 1 for progression-free survival

Characteristic	Loop 1	Loop 2
Inconsistency estimate (ABC)	0.09	–0.67
95% CI	–0.43 to 0.61	–1.45 to 0.11

Loop 1: A = everolimus, B = axitinib and C = sorafenib; Loop 2: A = everolimus, B = sorafenib and C = sunitinib.

Summary of mixed-treatment comparison model characteristics for overall survival analyses

	OS	
Characteristic	Primary analysis	SA
Total residual deviance	3	13
Number of data points	3	7

Results of the assessments for inconsistency in the data loops in mixed-treatment comparison sensitivity analysis for overall survival

Characteristic	Loop
Inconsistency estimate (ABC)	0.74
95% CI	0.15 to 1.33

Loop: A = everolimus, B = axitinib and C = sorafenib.

Appendix 6 Properties of health-related quality of life scales

Scale	Description	Range	Reported by
[a]EQ-5D-5L Index score[134,168]	Generic health scale	0–1	AXIS[43,67]
	Mobility, self-care, usual activities, pain/discomfort and anxiety/depression		CheckMate 025[54,68]
			METEOR[57,69]
EQ-VAS	Generic health scale	0–100	AXIS[43,67]
	20-cm visual scale from worst to best health imaginable		CheckMate 025[54,68]
			METEOR[57,69]
FKSI-19	Kidney cancer scale	0–76	METEOR[57,69]
	Disease-related symptoms (severity and interference with ADL), treatment side effects, function/well-being		
FKSI-15	Kidney cancer scale	0–60	AXIS[43,67]
	Shortened version of FKSI 19		
FKSI-DRS subscale[169]	Subscale assessing nine key symptoms: lack of energy, pain, weight loss, bone pain, fatigue, dyspnoea, cough, fevers, and haematuria	0–36	AXIS[43,67]
			CheckMate 025[54,68]
			RECORD-1[64,70]
EORTC QLQ-C30[170]	Cancer measure	0–100	RECORD-1[64,70]
Global health status and physical functioning	Single-item global health rating and multi-item physical functioning subscale of the full cancer quality-of-life measure		

ADL, activities of daily living.
a EQ-5D-5L codes were converted to a single index value normalised across all the patients using the UK algorithm. Index values range from 0 to 1. Higher scores on all scales indicate better health.

Appendix 7 Results of subgroup analyses

Progression-free survival subgroup analysis by Memorial Sloan Kettering Cancer Center baseline score; hazard ratio < 1 favours left-hand treatment

	HR (95% CrI)		
Comparison	Favourable	Intermediate	Poor
Cabozantinib vs. everolimus	0.51 (0.38 to 0.69)	0.47 (0.34 to 0.64)	0.72 (0.42 to 1.16)
Everolimus vs. BSC	0.30 (0.19 to 0.50)	0.32 (0.23 to 0.45)	0.42 (0.22 to 0.87)
Cabozantinib vs. BSC	0.15 (0.09 to 0.28)	0.15 (0.09 to 0.24)	0.28 (0.13 to 0.71)

Cells highlighted in green indicate statistically significant results.

Overall survival subgroup analysis by Memorial Sloan Kettering Cancer Center baseline score; hazard ratio < 1 favours left-hand treatment

	HR (95% CrI)		
Comparison	Favourable	Intermediate	Poor
Cabozantinib vs. everolimus	0.67 (0.46 to 0.95)	0.68 (0.48 to 0.94)	0.67 (0.39 to 1.07)
Nivolumab vs. everolimus	0.91 (0.59 to 1.33)	0.77 (0.58 to 0.99)	0.48 (0.30 to 0.73)
Nivolumab vs. cabozantinib	1.41 (0.78 to 2.34)	1.17 (0.74 to 1.74)	0.77 (0.37 to 1.42)

Cells highlighted in green indicate statistically significant results.

Progression-free survival subgroup analysis by number of prior therapies; hazard ratio < 1 favours left-hand treatment

	HR (95% CrI)	
Comparison	1 prior TKI	≥ 2 prior TKIs
Cabozantinib vs. everolimus	0.52 (0.41 to 0.66)	0.52 (0.35 to 0.74)
Everolimus vs. BSC	0.32 (0.24 to 0.43)	0.31 (0.19 to 0.54)
Cabozantinib vs. BSC	0.16 (0.11 to 0.24)	0.15 (0.08 to 0.31)

Cells highlighted in green indicate statistically significant results.

Overall survival subgroup analysis by number of prior therapies; hazard ratio < 1 favours left-hand treatment

Comparison	HR (95% CrI)	
	1 prior TKI	≥ 2 prior TKIs
Cabozantinib vs. everolimus	0.65 (0.50 to 0.85)	0.74 (0.48 to 1.10)
Nivolumab vs. everolimus	0.71 (0.56 to 0.90)	0.91 (0.61 to 1.30)
Cabozantinib vs. nivolumab	0.90 (0.64 to 1.30)	0.79 (0.47 to 1.43)

Cells highlighted in green indicate statistically significant results.

Appendix 8 Data abstraction tables

Clinical literature

AXIS

AXIS		Publication source
		Motzer 2013[43]
		Rini 2011[66]
		Rini 2015[148]
		Cella 2013[67]
		Escudier 2014[147]
		Motzer 2012[146]
Design		
Study design	Multicentre Phase III open-label RCT	Rini 2011[66]
Number of centres and country/countries	175 sites in 22 countries (Australia, Austria, Brazil, Canada, China, France, Germany, Greece, India, Ireland, Italy, Japan, South Korea, Poland, Russia, Singapore, Slovakia, Spain, Sweden, Taiwan, the UK and the USA)	Rini 2011[66]
Recruitment dates	15 September 2008 to 23 July 2010	Rini 2011[66]
Length of follow-up	Study start date: September 2008 Data cut-off point: July 2010 Completion date: February 2016	Rini 2011[66]
Source of funding	Pfizer	Rini 2011[66]
Eligibility criteria (inclusion and exclusion)	Inclusion: patients aged ≥ 18 years with histologically confirmed RCC, clear-cell component, measurable disease by RECIST; previous systemic first-line regimen with a sunitinib-based, bevacizumab plus interferon alfa-, temsirolimus- or cytokine-based regimen; ≥ 2 weeks since end of previous systemic treatment (≥ 4 weeks for bevacizumab plus interferon alfa); ECOG performance status of 0 or 1; life expectancy of ≥ 12 weeks; and adequate renal, hepatic and haematological organ function Exclusion: history of malignancy other than RCC; present use or anticipated need for cytochrome P450 drugs; known HIV or acquired immunodeficiency syndrome-related disease; CNS metastasis; uncontrolled hypertension; myocardial infarction, uncontrolled angina, congestive heart failure, or cerebrovascular accident within previous 12 months; and deep-vein thrombosis or pulmonary embolism within previous 6 months	Rini 2011[66]

AXIS			Publication source
Participants and treatment arms	***Intervention: axitinib***	***Comparator: sorafenib***	***Publication, data cut-off point (month, year)***
Intervention, method of delivery, dose and frequency	Orally at a starting dose of 5 mg twice daily and increased to 7 mg twice daily after 2 weeks and finally increased to 10 mg twice daily for patients without grade 2 or higher AEs and with blood pressure not higher than 150/90 mmHg Dose could be reduced to 3 mg twice daily or 2 mg twice daily if needed	Orally at a starting dose of 400 mg twice daily, which could be decreased to 400 mg once daily, and then to 400 mg every other day if dose reduction was needed due to toxic effects	Rini 2011[66]
Concomitant medication(s) or therapies	NR	NR	
Crossover or post-study interventions allowed (including number of patients)	Crossover was not allowed; no details on post-study medications reported	Crossover was not allowed; no details on post-study medications reported	Motzer 2013[43]
Number of cycles	NR	NR	
At least one dose reduction, *n* (%)	110 (31)	185 (52)	
Treatment duration (and the data cut-off points for each publication for the study)	Median duration of treatment was 8.2 months (range < 0.1–33.4 months)	Median duration of treatment was 5.2 months (range 0.2–34.1 months) with sorafenib	Motzer 2013[43] (November 2011)
Number randomised	361	362	Rini 2011[66]
Number who received study medication	359	355	Rini 2011[66]
Number withdrawn/ discontinued and reasons			Motzer 2013[43]
Total	318	325	
Disease progression/ relapse	240	226	
AEs	27	45	
Death	17	17	
Refusal of treatment for reason other than AEs	13	12	
Protocol violations	4	3	
Lost to follow-up	1	3	
Global deterioration in health	12	8	
Other reasons	4	11	
Disease stage and/or metastatic disease	Metastatic RCC	Metastatic RCC	Rini 2011[66]

AXIS			Publication source
Previous systemic therapy treatments, *n* (%)			Rini 2011[66]
Sunitinib	194 (54)	195 (54)	
Cytokines	126 (35)	125 (35)	
Bevacizumab	29 (8)	30 (8)	
Temsirolimus	12 (3)	12 (3)	
Age (years): median (range)	61 (20–82)	61 (22–80)	Rini 2011[66]
Ethnicity, *n* (%)			Rini 2011[66]
White	278 (77)	269 (74)	
Black	1 (< 1)	4 (1)	
Asian	77 (21)	81 (22)	
Other	5 (1)	8 (2)	
Male, *n* (%)	265 (73)	258 (71)	Rini 2011[66]
Performance status, *n* (%)			Rini 2011[66]
ECOG score of 0	195 (54)	200 (55)	
ECOG score of 1	162 (45)	160 (44)	
ECOG score of > 1	1 (< 1)	0	
MSKCC risk group			
Favourable	100 (28)	101 (28)	
Intermediate	134 (37)	130 (36)	
Poor	118 (33)	120 (33)	
N/A	9 (2)	11 (3)	
Heng risk factors			
Favourable	66 (18)	79 (22)	
Intermediate	236 (65)	225 (62)	
Poor	37 (10)	34 (9)	
N/A	22 (6)	24 (7)	
Reported subgroups	• OS by previous treatment, ECOG performance status (1 vs. 0), MSKCC risk score • PFS by previous treatment, ECOG performance status, MSKCC risk score, Heng risk score, ethnicity, gender, age and region		Motzer 2013[43]/ Rini 2011[66]
Reported outcomes			
Primary outcome	• PFS [time from randomisation to either first documentation of RECIST-defined disease progression (per independent radiology review of images) or death due to any cause, whichever came first]		Rini 2011[66]
Secondary outcomes	• OS (defined as the duration from assignment to study treatment to death) • ORR (confirmed CR or confirmed PR according to RECIST criteria • DOR (time in months from the first documentation of objective tumour response to objective tumour progression or death due to any cause) • Time to deterioration • Symptom deterioration using FKSI-15 and FKSI-DRS • AEs		Rini 2011[66] Motzer 2013[43] Cella 2013[67]

AXIS		Publication source
Outcomes and time points with data reported for subgroups of prior baseline therapies	• OS by previous treatment at data cut-off point 1 November 2011 • PFS by previous treatment at data cut-off point 31 August 2010	Motzer 2013[43] Rini 2011[66]
Outcomes and time points with data reported for subgroups of baseline prognostic scores (e.g. ECOG, MSKCC)	• OS by ECOG performance status (1 vs. 0), MSKCC risk group at data cut-off point 1 November 2011 • PFS by ECOG performance status, MSKCC risk score, and Heng risk score at data cut-off point 31 August 2010	Motzer 2013[43] Rini 2011[66]

Results	***Axitinib***	***Sorafenib***	***Publication and data cut-off point***
PFS			
HR (95% CI)	0.656 (0.552 to 0.779); $p < 0.0001$		Motzer 2013[43] (November 2011)
HR (95% CI) for subgroups based on prior therapy:			Motzer 2013[43] (November 2011)
Sunitinib	0.719 (0.572 to 0.903); $p = 0.0022$		
Cytokines	0.505 (0.373 to 0.684); $p < 0.0001$		
Bevacizumab + interferon alfa	0.815 (0.429 to 1.550); $p = 0.2656$		
Temsirolimus	1.210 (0.433 to 3.382); $p = 0.6342$		
PFS, median (95% CI) months	8.3 (6.7 to 9.2)	5.7 (4.7 to 6.5)	Motzer 2013[43] (November 2011)
PFS, median (95% CI), months for subgroups based on prior therapy			Motzer 2013[43] (November 2011)
Sunitinib	6.5 (5.7 to 7.9)	4.4 (2.9 to 4.7)	
Cytokines	12.2 (10.2 to 15.5)	8.2 (6.6 to 9.5)	
Bevacizumab + interferon alfa	8.3 (2.8 to 10.5)	4.5 (3.0 to 6.5)	
Temsirolimus	2.6 (1.5 to 17.1)	5.7 (2.6 to 8.3)	
Number of progression events, *n* (%)	NR	NR	
Overall survival			
HR (95% CI)	0.969 (0.800 to 1.174); $p = 0.3744$		Motzer 2013[43] (November 2011)
HR (95% CI) for subgroups based on prior therapy:			Motzer 2013[43] (November 2011)
Sunitinib	0.997 (0.782 to 1.270); $p = 0.4902$		
Cytokines	0.813 (0.555 to 1.191); $p = 0.1435$		
Bevacizumab + interferon alfa	1.825 (0.942 to 3.535); $p = 0.9648$		
Temsirolimus	0.459 (0.165 to 1.278); $p = 0.0638$		
Number of deaths, *n* (%)	211	214	Motzer 2013[43] (November 2011)
Median OS, months (95% CI)	20.1 (16.7 to 23.4)	19.2 (17.5 to 22.3)	Motzer 2013[43] (November 2011)

AXIS			Publication source
Median OS, months, for subgroup based on prior therapy			Motzer 2012[149] (November 2011)
Sunitinib	15.2	16.5	
Cytokine	29.4	27.8	
Bevacizumab + interferon alfa	14.7	19.8	
Temsirolimus	14.0	8.5	
Number of deaths, *n* (%) for subgroups based on prior therapy	NR	NR	
Response			
ORR, *n* (%)	82 (23)	45 (12)	Motzer 2013[43] (November 2011)
Independent review	70 (19)	34 (9)	Rini 2011[66] (August 2010)
CR, rate *n* (%)			Rini 2011[66] (August 2010)
Independent review	0	0	
Investigator assessment	0	1 (< 1)	
PR rate, *n* (%)			Rini 2011[66] (August 2010)
Independent review	70(19)	34 (9)	
Investigator assessment	70 (19)	39 (11)	
Stable disease ≥ 20 weeks, *n* (%) Independent review	96 (27)	77 (21)	Rini 2011[66] (August 2010)
Stable disease < 20 weeks, *n* (%) Investigator assessment	84 (23)	120 (33)	
Progressive disease, *n* (%)			Rini 2011[66] (August 2010)
Independent review	78 (22)	76 (21)	
Investigator assessment	60 (17)	66 (18)	
Time to response, months mean ± SD [median (range)]	NR	NR	
Duration of response, median, months (95% CI)	11 (7.4 to not estimable)	10.6 (8.8 to 11.5)	Rini 2011[66] (August 2010)
Other measures of response	NR	NR	

AXIS			Publication source
HRQoL			
Completion rates for all PRO			Cella 2013[67]
Baseline	> 85%	> 85%	
Cycle 8	52.1%	40.1%	
EQ-5D score			Cella 2013[67]
End of treatment, mean	0.71	0.69	
MD (95% CI) in EQ-5D score	0.02 (–0.01 to 0.05); $p = 0.1903$ (treatment by time $p = 0.8048$)		Cella 2013[67]
EQ-5D VAS			Cella 2013[67]
End of treatment, mean	68.11	68.64	
MD (95% CI) in EQ-5D VAS score	–0.53 (–2.77 to 1.72); $p = 0.6454$ (treatment by time $p = 0.1799$)		Cella 2013[67]
FKSI-15			Motzer 2013[43] (November 2011)
Baseline score, mean ± SD	43.2 ± 8.4	43.3 ± 8.2	
End of treatment score, mean	42.21	41.86	Cella 2013[67]
MD (95% CI) in FKSI-15 score	0.35 (–0.63 to 1.34); $p = 0.4833$ (treatment by time $p = 0.3943$)		Cella 2013[67]
FKSI-DRS			Motzer 2013[43] (November 2011)
Baseline score, mean ± SD	28.9 ± 5.2	29.0 ± 5.2	
End of treatment, mean	28.56	28.44	Cella 2013[67]
MD (95% CI) in FKSI-DRS score	0.12 (–0.45 to 0.69); $p = 0.6746$ (treatment by time $p = 0.8024$)		Cella 2013[67]
AE grade ≥ 3, *n* (%)			
n in safety analysis	359	355	Motzer 2013[43] (November 2011)
Total AE grade ≥ 3	NR	NR	
Total AEs (any grade)	NR	NR	
Stomatitis	5 (1)	1 (< 0.5)	
Rash	1 (< 0.5)	13 (4)	
Fatigue	37 (10)	14 (4)	
Asthenia	15 (4)	8 (2)	
Diarrhoea	40 (11)	27 (8)	
Anorexia	15 (4)	7 (2)	
Nausea	6 (2)	3 (1)	
Vomiting	5 (1)	0	
Cough	NR	NR	
Dry skin	0	0	

AXIS			Publication source
Infection	NR	NR	
Pneumonitis	NR	NR	
Dyspnoea	NR	NR	
Anaemia	NR	NR	
Hypertension	60 (17)	43 (12)	
Dysphonia	0	0	
Hand–foot syndrome	20 (6)	61 (17)	
Hypothyroidism	1 (< 0.5)	0	
Weight decreased	12 (3)	9 (3)	
Mucosal inflammation	5 (1)	3 (1)	
Constipation	1 (< 0.5)	1 (< 0.5)	
Proteinuria	11 (3)	4 (1)	
Dysgeusia	0	0	
Headache	3 (1)	0	
Arthralgia	3 (1)	1 (< 0.5)	
Alopecia	0	0	
Pruritus	0	0	
Pain in extremity	1 (< 0.5)	3 (1)	
Erythaema	0	1 (< 0.5)	

CNS, central nervous system; CR, complete response; DOR, duration of response; HIV, human immunodeficiency virus; N/A, not available; NR, not reported; PR, partial response; PRO, patient reported outcomes; PS, performance status.

Calvani et al.[58]

Calvani *et al.*[58]	
Design	
Study design	Retrospective sequencing study
Number of centres and country/countries	Three oncology centres in Italy
Cohort recruitment	NR
Recruitment dates	Patients treated between January 2006 and October 2010
Length of follow-up	NR
Source of funding	NR
Eligibility criteria (inclusion and exclusion)	Patients with stage IV RCC who had experienced disease progression or unacceptable toxicity after receiving either sorafenib or sunitinib as first TKI and then switched to the other reciprocal agent as second TKI. Elapsed time between the TKIs was ≤ 2 months

Calvani *et al.*[58]			
Participants and treatment arms	***Intervention: sunitinib***	***Comparator: sorafenib***	
Intervention, method of delivery, dose and frequency	50 mg daily, 4 weeks on and 2 weeks off Dose reduction, delays or discontinuation were determined independently by each investigator	400 mg twice daily Dose reduction, delays or discontinuation were determined independently by each investigator	
Concomitant medication(s) or therapies	NR	NR	
Crossover or post-study interventions allowed (including number of patients)	NR	NR	
Number of cycles, dose reductions	NR	NR	
Treatment duration (and the data cut-off points for each publication for the study)	NR	NR	
Number randomised	N/A	N/A	
Number who received study medication	15	18	
Withdrawn/discontinued *n* (%) and reasons			
Lost to follow-up	3 (20)	3 (17)	
Disease stage and/or metastatic disease, *n* (%)	Stage IV (100)	Stage IV (100)	
≥ 2 metastatic sites, *n* (%)	9 (60)	5 (25)	
Previous systemic therapy treatments, *n* (%)			
First-line treatment TKI	11 (73)	14 (78)	$p = 1$
First-line treatment cytokines	4 (27)	4 (22)	
Previous TKI	Sorafenib 15 (100)	Sunitinib 18 (100)	
Age (years): mean ± SD (range)	70 (50–74)	61 (46–73)	$p = 0.0429$
Ethnicity, *n* (%)	NR	NR	
Male, *n* (%)	12 (80)	11 (61)	$p = 0.28$
Performance status			$p = 0.35$
ECOG score of			
0–1	14 (93)	14 (78)	$p = 1$
2	1 (7)	4 (22)	
MSKCC			
Good	3 (20)	4 (22)	
Intermediate	11 (73)	14 (78)	
Poor	1 (7)	0 (0)	
Reported subgroups	Age: elderly (≥ 65 years old) and young adult (< 65 years old)		
Reported outcomes			
Primary outcome	PFS on first and second TKI. PFS on second TKI defined as time from the start of treatment with the targeted agent to disease progression, death or discontinuation due to intolerance. Patients who did not experience progression or were lost on follow-up under the second TKI were censored. Total PFS was defined as the sum of PFS on first and second TKI excluding any elapsed time between the two treatment periods. Disease progression was assessed according to RECIST		
Secondary outcomes	Overall survival was define as the time from administration of first TKI to death from any cause. Patients lost to follow-up were censored		

Calvani *et al.*[58]		
Results	***Intervention: sunitinib***	***Comparator: sorafenib***
PFS		
HR (95% CI)	0.46 (0.16 to 0.95); $p = 0.0377$	
HR (95% CI) for subgroups based on prior therapy	NR	
PFS median (range) months	11	3
PFS mean ± SD	NR	NR
[median (range)], months for subgroups based on prior therapy		
Number of progression events, *n* (%)	NR	NR

N/A, not applicable; NR, not reported.

CheckMate 025

CheckMate 025		Publication source
		Motzer 2015[54]
		Motzer 2015 supplement[54]
		Motzer 2016[149]
		Cella 2016[150]
		Cella 2016[68]
Design		
Study design	Randomised, open-label, Phase III trial	Motzer 2015[54]
Number of centres and country/countries	146 sites in 24 countries (Argentina, Australia, Austria, Belgium, Brazil, Canada, Czech Republic, Denmark, Finland, France, Germany, Greece, Ireland, Israel, Italy, Japan, Norway, Poland, Romania, Russia, Spain, Sweden, the UK and the USA)	Motzer 2015[54] Motzer 2015[54] supplement
Recruitment dates	October 2012 to March 2014	Motzer 2015[54]
Length of follow-up	October 2012 to June 2015 (minimum follow-up period was 14 months)	Motzer 2015[54]
Source of funding	Bristol-Myers Squibb	Motzer 2015[54]
Eligibility criteria (inclusion and exclusion)	Inclusion: aged ≥ 18 years with histological confirmation of amRCC with a clear-cell component and measurable disease according to RECIST version 1.1; received 1 or 2 previous regimens of antiangiogenic therapy but no more than 3 total previous regimens of systemic therapy, including cytokines and cytotoxic chemotherapy drugs; disease progression during or after the last treatment regimen and within 6 months before study enrolment; and Karnofsky performance status of ≥ 70 at the time of study entry Exclusion: patients with metastasis to the central nervous system, previous treatment with an mTOR inhibitor, or a condition requiring treatment with glucocorticoids (equivalent to > 10 mg of prednisone daily)	Motzer 2015[54]

CheckMate 025			Publication source
Participants and treatment arms	***Intervention: nivolumab***	***Comparator: everolimus***	
Intervention, method of delivery, dose and frequency	Nivolumab at a dose of 3 mg per kilogram of body weight as a 60-minute intravenous infusion every 2 weeks	Everolimus, orally as a daily dose of 10 mg	Motzer 2015[54]
Concomitant medication(s) or therapies	NR		
Crossover or post-study interventions allowed (including number of patients)			Motzer 2015[54]
Subsequent systemic therapy, *n* (%)	227 (55)	260 (63)	
Everolimus	105 (26)	N/A	
Axitinib	99 (24)	149 (36)	
Pazopanib	37 (9)	64 (16)	
Sorafenib	0	38 (9)	
Anti-PD-1	0	7 (2)	
Number of cycles, dose reductions	Dose modifications were not permitted for nivolumab but were permitted for everolimus; 102/397 everolimus patients had at least one dose reduction		Motzer 2015[54]
Median treatment duration (range)	5.5 months (< 0.1–29.6)	3.7 months (0.2–25.7)	Motzer 2015[54]
Number randomised	410	411	Motzer 2015[54]
Number who received study medication	406	397	Motzer 2015[54]
	207 had dose delays during study	262 had dose delays during study	
Number withdrawn/discontinued and reasons			Motzer 2015[54] supplement
Did not receive study drug	4	14	
Discontinued intervention:	339	369	
Disease progression	285	273	
Study drug toxicity	35	53	
AE unrelated to study drug	9	14	
Request to discontinue treatment	5	14	
Other	5	11	
Disease stage and/or metastatic disease	amRCC	amRCC	Motzer 2015[54]
Previous systemic therapy treatments, *n* (%)			Motzer 2015[54]
Sunitinib	246 (60)	242 (59)	
Pazopanib	119 (29)	131 (32)	
Axitinib	51 (12)	50 (12)	
Number of previous antiangiogenic regimens, *n* (%)			
1	294 (72)	297 (72)	
2	116 (28)	114 (28)	
Age (years): median (range)	62 (23–88)	62 (18–86)	Motzer 2015[54]

CheckMate 025			Publication source
Ethnicity, *n* (%)			Motzer 2015[54]
White	353 (86)	367 (89)	
Asian	42 (10)	32 (8)	
Black	1 (< 1)	4 (1)	
Other	14 (3)	8 (2)	
Male, *n* (%)	315 (77%)	304 (74%)	Motzer 2015[54]
Performance status			
MSKCC			
Favourable	145 (35)	148 (36)	Motzer 2015[54]
Intermediate	201 (49)	203 (49)	
Poor	64 (16)	60 (15)	
Karnofsky score			
< 70	2 (< 1)	1 (< 1)	
70	22 (5)	30 (7)	
80	110 (27)	116 (28)	
90	150 (37)	130 (32)	
100	126 (31)	134 (33)	
Reported subgroups	OS by region, MSKCC risk score, number of previous anti-angiogenic therapy, PD-L1 status, age and sex		Motzer 2015[54]
Reported outcomes			
Primary outcome	OS (defined as the time from randomisation to the date of death)		Motzer 2015[54]
Secondary outcomes	ORR, PFS, the association between OS and tumour expression of PD-L1, the incidence of AEs, HRQoL (FKSI-DRS – baseline then 4 weekly to week 104)		Motzer 2015[54]
Outcomes and time points with data reported for subgroups of prior baseline therapies	Data cut-off point at June 2015 OS by number of previous anti-angiogenic therapy		Motzer 2015[54]
Outcomes and time points with data reported for subgroups of baseline prognostic scores (e.g. ECOG, MSKCC)	Data cut-off point at June 2015 OS by MSKCC risk score		Motzer 2015[54]
Results	***Nivolumab***	***Everolimus***	***Publication and data cut-off point date (month, year)***
PFS			
HR (95% CI)	0.88 (0.75 to 1.03); $p = 0.11$		Motzer 2015[54]
HR (95% CI) for subgroups based on prior therapy	NR		
PFS, median (95% CI) months	4.6 (3.7 to 5.4)	4.4 (3.7 to 5.5)	Motzer 2015[54]
PFS mean ± SD [median (range)], months for subgroups based on prior therapy	NR	NR	
Number of progression events, *n* (%)	318 (78)	322 (78)	Motzer 2015[54]
At 6 months	145 (35%)	129 (31%)	

CheckMate 025			Publication source
Overall survival			
HR, (98.5% CI)	0.73 (0.57 to 0.93); $p = 0.002$		Motzer 2015[54]
OS, median (95% CI), months	25.0 (21.8 to not estimable)	19.6 (17.6 to 23.1)	Motzer 2015[54]
OS Median (95% CI) by KPS, %			Motzer 2016[149]
90–100 ($n = 540$)	Not estimable (26.7 to not estimable)	29.0 (24.3 to not estimable)	
< 90 ($n = 281$)	18.1 (14.3 to 22.2)	10.1 (7.9 to 12.8)	
OS median (95% CI) by Heng risk group			Motzer 2016[149]
Favourable ($n = 125$)	Not estimable	29.0 (24.7 to not estimable)	
Intermediate ($n = 483$)	Not estimable (21.4 to not estimable)	19.9 (17.7 to 26.2)	
Poor ($n = 179$)	15.3 (10.6 to 20.4)	8.4 (5.9 to 11.4)	
OS median (95% CI) by prior antiangiogenic			Motzer 2016[149]
1 prior ($n = 629$)	23.6 (20.8 to not estimable)	19.9 (17.7 to 24.7)	
2 prior ($n = 189$)	Not estimable (18.1 to not estimable)	18.4 (14.0 to not estimable)	
HR (95% CI) for subgroups based on prior therapy			Motzer 2015[54]
1 previous antiangiogenic regimen	0.71 (0.56 to 0.90)		
2 previous antiangiogenic regimens	0.89 (0.61 to 1.29)		
HR (95% CI) for subgroups based on performance status			Motzer 2015[54]
MSKCC favourable	0.89 (0.59 to 1.32)		
MSKCC intermediate	0.76 (0.58 to 0.99)		
MSKCC poor	0.47 (0.30 to 0.73)		
Number of deaths, *n* (%)	183 (45)	215 (52)	Motzer 2015[54]
Number of deaths, *n* (%) for subgroups based on prior therapy:			Motzer 2015[54]
1 antiangiogenic regimen	166/294	139/297	
2 antiangiogenic regimens	61/116	57/114	
Response			
ORR, *n* (%)	103 (25)	22 (5)	Motzer 2015[54]
	OR 5.98 (95% CI 3.68 to 9.72); $p < 0.001$		
ORR % (95% CI) by KPS %			Motzer 2016[149]
90–100 ($n = 540$)	26.1 (21.0 to 31.7)	6.8 (4.1 to 10.6)	
< 90 ($n = 281$)	23.1 (16.3 to 31.2)	2.7 (0.7 to 6.8)	
ORR % (95% CI) by Heng risk group			Motzer 2016[149]
Favourable ($n = 125$)	23.6 (13.2 to 37.0)	7.1 (2.4 to 15.9)	
Intermediate ($n = 483$)	24.4 (19.1 to 30.3)	5.0 (2.6 to 8.5)	
Poor ($n = 179$)	30.2 (21.3 to 40.4)	4.8 (1.3 to 11.9)	

CheckMate 025			Publication source
ORR% (95% CI) by prior anti-angiogenic			Motzer 2016[149]
1 prior ($n = 629$)	24.3 (19.7 to 29.4)	5.4 (3.2 to 8.6)	
2 prior ($n = 189$)	27.8 (18.9 to 38.2)	5.1 (1.7 to 11.4)	
Complete response, rate *n* (%)	4 (1)	2 (< 1)	Motzer 2015[54] supplement
PR rate, *n* (%)	99 (24)	20 (5)	Motzer 2015[54] supplement
Stable disease, *n* (%)	141 (34)	227 (55)	Motzer 2015[54] supplement
Time to response, median (range), months	3.5 (1.4–24.8) ($n = 103$)	3.7 (1.5–11.2) ($n = 22$)	Motzer 2015[54]
Duration of response, median (range), months	12.0 (0–27.6)	12.0 (0–22.2)	Motzer 2015[54]
Progressive disease, *n* (%)	143 (35)	114 (28)	Motzer 2015[54] supplement
Ongoing response (among patients with a treatment response), *n* (%)	10 (45)	49 (48)	Motzer 2015[54]
Ongoing response for 12 months or longer (among patients with a treatment response), *n* (%)	6 (27)	32 (31)	Motzer 2015[54]
HRQoL			
Completion rates *n* (%)			Motzer 2015[54] supplement
First year (week 1 – week 52)	≥ 80	≥ 80	
> 1 year lowest rate (week 53- week 104)	71	60	
FKSI-DRS, baseline median (range) score	31.0	31.0	Motzer 2015[54] supplement
			Cella 2016[68]
FKSI-DRS baseline score, mean (SD)	30.2 (4.4)	30.1 (4.8)	
EQ-5D index baseline score, mean (SD)	0.78 (0.24)	0.78 (0.21)	
EQ-5D VAS baseline score, mean (SD)	73.3 (18.5)	72.5 (18.7)	
	361/406 treated had baseline FKSI-DRS (89%)	343/397 treated had baseline FKSI-DRS (86%)	
	361/406 treated had baseline EQ-5D (89%)	344/397 treated had baseline EQ-5D (87%)	
FKSI-DRS, change from baseline Median (range) at week 104	2.0 (–1.0 to 16.0) [$n = 20$ (77%)]	–2.0 (–7.0 to 15.0) [$n = 9$ (90%)]	Motzer 2015[54] supplement
	$p < 0.05$ vs. everolimus		
Least squares mean (SE) with repeated measures mixed-effects model to week 84 (results not given to week 104)	–1.8 (0.2); $p < 0.0001$	–0.2 (0.2); $p = 0.44$	Cella 2016[68]
Difference from everolimus in mean change to week 84, mean (95% CI)	1.7 (1.2 to 2.1); $p < 0.0001$		

CheckMate 025			Publication source
FKSI-DRS, *n* (%) with ≥ 2-point increase over course of study	200/361 (55)	126/343 (37) (difference $p < 0.001$)	Cella 2016[68]
FKSI-DRS median time to ≥ 2-point improvement, months (95% CI)	4.7 (3.7 to 7.5)	Not reached (not estimable)	Cella 2016[68]
EQ-5D, difference in mean change from baseline to end point. Mixed-effects repeated measures model (95% CI)	EQ-5D utility index 0.04 (0.02 to 0.07); $p = 0.0003$ EQ-5D VAS 5.7 (3.8 to 7.7); $p < 0.0001$		Cella 2016[68]
EQ-5D VAS *n* (%) with 7-point improvement	192/360 (53) Difference vs. everolimus $p = 0.0001$)	134/343 (39)	Cella 2016[68]
EQ-5D median time to 7 + point improvement, months (95% CI)	6.5 (3.9 to 12.2) Difference HR 1.37, 1.10 to 1.71; $p = 0.0054$	23.1 (15.4 to not estimable)	Cella 2016[68]
AEs grade ≥ 3, *n* (%)			
n in safety analysis	406	397	
Any grade	319 (79)	349 (88)	Motzer 2015[54]
Grade 3 or 4	76 (19)	145 (37)	Motzer 2015[54]
Deaths attributed to study drug toxic effects	0	2	Motzer 2015[54]
Fatigue	10 (2)	11 (3)	Motzer 2015[54]
Nausea	1 (< 1)	3 (1)	Motzer 2015[54]
Pruritus	0	0	Motzer 2015[54]
Diarrhoea	5 (1)	5 (1)	Motzer 2015[54]
Decreased appetite	2 (< 1)	4 (1)	Motzer 2015[54]
Rash	2 (< 1)	3 (1)	Motzer 2015[54]
Cough	0	0	Motzer 2015[54]
Anaemia	7 (2)	31 (8)	Motzer 2015[54]
Dyspnoea	3 (1)	2 (1)	Motzer 2015[54]
Peripheral oedema	0	2 (1)	Motzer 2015[54]
Pneumonitis	6 (1)	11 (3)	Motzer 2015[54]
Mucosal inflammation	0	12 (3)	Motzer 2015[54]
Dysgeusia	0	0	Motzer 2015[54]
Hyperglycaemia	5 (1)	15 (4)	Motzer 2015[54]
Stomatitis	0	17 (4)	Motzer 2015[54]
Hypertriglyceridaemia	0	20 (5)	Motzer 2015[54]
Epistaxis	0	0	Motzer 2015[54]

KPS, Karnofsky performance status; N/A, not applicable; NR, not reported; PD-1, programmed cell death protein 1; PD-L1, programmed death receptor ligand 1; PR, partial response.

ESPN

ESPN			Publication source
			Tannir 2016[55]
			Tannir 2014[151]
Design			
Study design	Randomised multicentre Phase II trial with crossover study phase		Tannir 2016[55]
Number of centres and country/countries	Four locations in the USA		
Recruitment dates	3 September 2010 to 19 November 2013		Tannir 2016[55]
Length of follow-up	September 2010 to final analysis on May 2014. Median follow-up of 23.6 months (95% CI 15.7 to 30.2)		Tannir 2016[55]
Source of funding	Novartis		Tannir 2016[55]
Eligibility criteria (inclusion and exclusion)	Inclusion: patients were > 18 years of age and had not received prior systemic therapy for advanced papillary, chromophobe, CDC, Xp11.2 translocation, unclassified RCC, or ccRCC with > 20% sarcomatoid features in their primary tumours; ECOG performance status of 0 or 1, measurable disease, and adequate organ and marrow function Exclusion: patients with untreated brain metastases, metabolic dysfunction and uncontrolled medical conditions		Tannir 2016[55]
Participants and treatment arms	***Intervention: sunitinib***	***Comparator: everolimus***	***Publication and data cut-off point (month, year)***
Intervention, method of delivery, dose and frequency	First line: everolimus 10 mg/day orally for 4 weeks on and 2 weeks off. Patients were treated until progressive disease, unacceptable AEs, or withdrawal of consent. At progressive disease, patients were allowed to receive the agent that they did not receive upfront Second line: sunitinib 50 mg/day orally	First line: sunitinib 50 mg/day orally for 4 weeks on and 2 weeks off. Patients were treated until progressive disease, unacceptable AEs, or withdrawal of consent. At progressive disease, patients were allowed to receive the agent that they did not receive upfront Second line: everolimus 10 mg/day orally	Tannir 2016[55]
Concomitant medication(s) or therapies	NR	NR	Tannir 2016[55]
Crossover or post-study interventions allowed	NR	NR	Tannir 2016[55]
Number of cycles, dose reductions, *n* (%)	Number of cycles: NR		Tannir 2016
Second-line sunitinib dose reductions	4 (19) [pre-crossover everolimus dose reductions 5 (14)]		
Second-line everolimus dose reductions	1 (4) [pre-crossover sunitinib dose reductions 13 (39)]		
Treatment duration (and the data cut-off points for each publication for the study)			Tannir 2016[55] (May 2014)
Median follow-up (95% CI)	23.6 months (15.7 to 30.2)	23.6 months (15.7 to 30.2)	
Treatment duration	NR	NR	
Number randomised	N/A	N/A	Tannir 2016[55]

ESPN			Publication source
Number who received study medication	21	23	Tannir 2016[55]
Number withdrawn/discontinued and reasons			Tannir 2016[55]
Total	19	19	
Disease progression	17	16	
AEs	1	1	
Death	1	NR	
Physician's decision	NR	1	
Patient withdrew consent	NR	1	
Disease stage and/or metastatic disease	Metastatic RCC	Metastatic RCC	Tannir 2016[55]
Previous systemic therapy treatments, *n* (%)	Patients had not received prior systemic therapy for advanced RCC	Patients had not received prior systemic therapy for advanced RCC	Tannir 2016[55]
Age (years): median (range)	58 (23–73)	60 (28–76)	Tannir 2016[55]
First line			
Ethnicity, *n* (%)			Tannir 2016[55]
First line:			
White	28	25	
Hispanic	3	5	
Black	2	3	
Male, *n* (%)			Tannir 2016[55]
First line	24	19	
Performance status			Tannir 2016[55]
First line			
ECOG score of 0	15	18	
ECOG score of 1	20	15	
MSKCC risk group			
Good	4	4	
Intermediate	29	29	
Poor	2	0	
IMDC risk group			
Good	4	3	
Intermediate	24	26	
Poor	7	4	
Reported subgroups	Histological RCC subtype		Tannir 2016[55]
Reported outcomes			
Primary outcome	PFS in first line (calculated from date of randomisation to the date of first documented progressive disease, or death from any cause)		Tannir 2016[55]
Secondary outcomes	ORR in first-line therapy, ORR in second-line therapy, OS, PFS in second-line therapy and safety		Tannir 2016[55]

ESPN			Publication source
Outcomes and time points with data reported for subgroups of prior baseline therapies	NR		
Outcomes and time points with data reported for subgroups of baseline prognostic scores (e.g. ECOG, MSKCC)	NR		
Results	***Intervention: sunitinib***	***Comparator: everolimus***	***Publication and data cut-off point (month, year)***
PFS			
HR (95% CI)	NR		
HR (95% CI) for subgroups based on prior therapy	NR		
PFS median (95% CI) months	1.8 (1.4 to 10.6); $p = 0.6$ vs. everolimus Figure 2B in manuscript is KM curve for second-line PFS	2.8 (1.4 to not available)	Tannir 2016[55] (May 2014)
PFS mean ± SD	NR	NR	
Number of progression events *n* (%)	16 patients	17 patients	Tannir 2016[55]
Response			
ORR, *n* (%)	2	2	Tannir 2016[55]
CR, rate, *n* (%)	NR	NR	
PR rate, *n* (%)	2	2	Tannir 2016[55]
Stable disease, *n* (%)	7	9	Tannir 2016[55]
Time to response, months	NR	NR	
Duration of response, months	NR	NR	
Progressive disease	10	9	Tannir 2016[55]
AEs grade ≥ 3, *n* (%) (note: AEs data are presented as combined first and second line)			
Total AEs grade ≥ 3	19 (54)	29 (88)	Tannir 2016[55]
Fatigue	2	13	Tannir 2016[55]
Diarrhoea	1	8	Tannir 2016[55]
Anorexia			Tannir 2016[55]
Nausea	1	4	Tannir 2016[55]
Anaemia	5	6	Tannir 2016[55]
Hypertension	0	9	Tannir 2016[55]
Neutropenia	0	7	Tannir 2016[55]
Thrombocytopenia	1	4	Tannir 2016[55]
Hyponatremia	0	4	Tannir 2016[55]

ccRCC, clear-cell renal cell carcinoma; CDC, collecting duct carcinoma; CR, complete response; IMDC, International Metastatic Renal Cell Carcinoma Database Consortium; N/A, not available; NR, not reported; PR, partial response; PS, performance status.

Iacovelli et al.[59]

Iacovelli *et al.*[59]		
Design		
Study design	Retrospective observational study	
Number of centres and country/countries	23 centres in Italy	
Recruitment dates	NR	
Length of follow-up	NR	
Source of funding	NR	
Eligibility criteria (inclusion and exclusion)	Inclusion: patients with clear-cell mRCC treated with everolimus or sorafenib as third-line therapy	
Participants and treatment arms	***Intervention: sorafenib***	***Comparator: everolimus***
Intervention, method of delivery, dose and frequency	Standard dose	
Concomitant medication(s) or therapies	NR	
Crossover or post-study interventions allowed (including number of patients)	NR	
Number of cycles, dose reductions	NR	
Treatment duration (and the data cut-off points for each publication for the study)	Treatments were administered until disease progression or until the patient developed unacceptable levels of toxicity	
Number randomised	N/A	
Number who received study medication, *n* (%)	90 (38.6)	143 (61.4)
Number withdrawn/discontinued and reasons	NR	NR
≥ 2 sites of metastases, *n* (%)	(86.2)	
Previous systemic therapy treatments, *n* (%)		
First-line sunitinib	(66)	
Sorafenib	(19)	
Bevacizumab +	(10)	
Interferon		
Other	(5)	
Second-line sunitinib	(31)	
Sorafenib	(33)	
Everolimus	(25)	
Temsirolimus	(10)	
Age (years): median (IQR)	63.2 (55.7–70.9)	
Ethnicity, *n* (%)	NR	
Male, *n* (%)	(73.8)	
Performance status, *n* (%)		
ECOG score		
0	(28.4)	
1	(52.6)	
2	(19.0)	

Iacovelli *et al.*[59]		
Heng		
Good	(22.7)	
Intermediate	(69.1)	
Poor	(8.2)	
Reported subgroups	NR	
Reported outcomes		
Outcomes	PFS was defined as the time from beginning of treatment to progression or death from any cause, whichever occurred first. The progression of disease was defined as a ≥ 20% increase of the long diameter according to the RECIST 1.0 criteria. Response assessment by CT or MRI scans was carried out according to local procedures every 8 to 12 weeks and assessed locally by a radiologist OS was defined as the time from start of third-line treatment to death or censored at last contact	
Results	***Intervention: sorafenib***	***Comparator: everolimus***
Overall survival		
HR (95% CI) univariate Cox regression	2.43 (1.65 to 3.63); $p < 0.001$	
HR (95% CI) multivariable Cox regression (adjusted for Heng prognostic criteria)	2.21 (1.47 to 3.31); $p < 0.001$	
Number of deaths, *n* (%)	NR	NR

IQR, interquartile range; N/A, not applicable; NR, not reported.

METEOR

METEOR		Publication source
		Choueiri 2016 appendix[57]
		Choueiri 2016[57]
		Choueiri 2015[145]
		Cella 2016[69]
		NICE TA10075[28]
Design		
Study design	Randomised, open-label, Phase III study	Choueiri 2016[57]
Number of centres and country/ countries	Patients enrolled at 173 hospital and outpatient clinics in 26 countries	Choueiri 2016[57]
Recruitment dates	Between 8 August 2013 and 24 November 2014	Choueiri 2016[57]
Length of follow-up	The median duration of follow-up for overall survival and safety was 18.7 months (IQR 16.1–21.1) in the cabozantinib group and 18.8 months (IQR 16.0–21.2) in the everolimus group. The median duration of follow-up for PFS was 11.4 months (IQR 8.8–13.7) in the cabozantinib group and 11.5 months (IQR 8.6–13.9) in the everolimus group	Choueiri 2016[57]

METEOR			Publication source
Source of funding	Exelixis, CA, USA		Choueiri 2016[57]
Eligibility criteria (inclusion and exclusion)	Patients aged ≥ 18 years with amRCC and a clear-cell histology, with measurable disease per RECIST (version 1.1), who had received at least one previous VEGFR TKI (there was no limit to the number of previous treatments), and had disease progression during or within 6 months of the most recent VEGFR TKI treatment and within 6 months before randomisation. Patients were required to have a Karnofsky performance status score of at least 70% and adequate organ function, based on standard laboratory tests including haematology, serum chemistry, lipids, coagulation, thyroid function and urinalysis. Patients with brain metastases were allowed provided these were stable and asymptomatic		Choueiri 2016[57]
	Patients with previous mTOR inhibitor therapy, including everolimus, were not eligible for the study, and nor were patients with uncontrolled hypertension or clinically significant cardiovascular, gastrointestinal, wound healing or infectious comorbidities		
Participants and treatment arms	***Intervention: cabozantinib***	***Comparator: everolimus***	***Publication***
Intervention, method of delivery, dose and frequency	Orally once a day at 60 mg	Orally once a day at 10 mg	Choueiri 2016[57]
Concomitant medication(s) or therapies	NR	NR	
Crossover	On study crossover between treatment groups was not permitted		Choueiri 2016[57]
Subsequent systemic anticancer treatment, *n* (%)	165 (50)	181 (55)	Choueiri 2016,[57] appendix
Axitinib	57 (17)	90 (27)	
Cabozantinib	0	7 (2)	
Pazopanib	5 (2)	22 (7)	
Sorafenib	9 (3)	31 (9)	
Sunitinib	17 (5)	33 (10)	
Everolimus	96 (29)	15 (5)	
Temsirolimus	6 (2)	4 (1)	
Bevacizumab	8 (2)	11 (3)	
Interleukin 2	0	4 (1)	
Interferon-alpha	5 (2)	7 (2)	
PD-1/PD-L1 targeting agents	15 (5)	19 (6)	
Chemotherapy	11 (3)	13 (4)	
Median daily dose, mg (IQR)	Cabozantinib could be dose reduced to 40 mg and then 0 mg	Everolimus could be dose reduced to 5 mg and then 2.5 mg	Choueiri 2016[57]
	43 (36–56)	9 (7–10)	
Dose reductions, *n* (%)	206 (62)	80 (25)	
	Patients were allowed to continue study treatment beyond radiographic progression at the discretion of the investigator		

METEOR			Publication source
Treatment duration, months (IQR)	8.3 (4.2 to 14.6)	4.4 (1.9 to 8.6)	Choueiri 2016[57]
Median duration (months)	7.6	4.4	Choueiri 2015[145]
Number randomised	330	328	Choueiri 2016[57]
Number who received study medication	331	322	Choueiri 2016[57]
Number withdrawn/discontinued and reasons			Choueiri 2016[57]
Discontinued cabozantinib	257	297	
Disease progression	159	190	
AEs	40	34	
Deaths not treatment related	2	NR	
Death treatment related	1	1	
Clinical deterioration	35	52	
Withdrew consent	8	13	
Other	15	8	
Disease stage and/or metastatic disease	Advanced or metastatic	Advanced or metastatic	
Previous systemic therapy treatments, *n* (%) (ITT population)			Choueiri 2016[57]
Prior TKI			
1	235 (71)	229 (70)	
≥ 2	95 (29)	99 (30)	
Sunitinib	210 (64)	205 (62)	
Pazopanib	144 (44)	136 (41)	
Axitinib	52 (16)	55 (17)	
Sorafenib	21 (6)	31 (9)	
Bevacizumab	5 (2)	11 (3)	
Interleukin 2	20 (6)	29 (9)	
Interferon-alpha	19 (6)	24 (7)	
Nivolumab	17 (5)	14 (4)	
Age (years): mean ± SD (range)	63 (32–86)	62 (31–84)	Choueiri 2015[145]
Ethnicity, *n* (%)			Choueiri 2016[57]
White	269 (82)	263 (80)	
Asian	21 (6)	26 (8)	
Black	6 (2)	3 (< 1)	
Other	19 (6)	13 (4)	
Not reported	15 (5)	22 (7)	
Missing data	0	1 (< 1)	
Male, *n* (%)	253 (77)	241 (73)	Choueiri 2016[57]
Performance status, *n* (%)			Choueiri 2016[57]
ECOG score			
0	226 (68)	217 (66)	
1	104 (32)	111 (34)	

METEOR			Publication source
MSKCC risk group			
Favourable	150 (45)	150 (46)	
Intermediate	139 (42)	135 (41)	
Poor	41 (12)	43 (13)	
Reported subgroups	Age, sex, race, MSKCC risk group, previous nephrectomy, ECOG status, diagnosis to randomisation, tumour MET status, number of organs with metastases, SoD, bone metastases, visceral metastases, visceral and bone metastases, number of previous VEGFR-TKIs, duration of first VEGFR-TKI, progression after start of most recent VEGFR-TKI, previous PD-1 or PD-L1 treatment, only previous VEGFR-TKI		Choueiri 2016[57]
Reported outcomes			
Primary outcome	PFS by independent radiology review in the first 375 randomised patients. PFS was defined as the time from randomisation to radiographic progression per RECIST or death from any cause		Choueiri 2016[57]
Secondary outcomes	OS was defined as the time from randomisation to death from any cause, and objective response per independent radiology review committee assessment, defined as the proportion of patients with a confirmed complete or partial response per RECIST, assessed in all randomly assigned patients Safety and tolerability were also assessed		Choueiri 2016[57]
Outcomes and time points with data reported for subgroups of prior baseline therapies	OS and PFS by number of previous VEGFR-TKIs (1 or ≥ 2)		Choueiri 2016[57]
Outcomes and time points with data reported for subgroups of baseline prognostic scores (e.g. ECOG, MSKCC)	OS and PFS by MSKCC risk group (favourable, intermediate or poor) and ECOG status (0 or 1)		Choueiri 2016[57]
Results *PFS*	***Intervention: cabozantinib***	***Comparator: everolimus***	***Publication and data cut-off point (month, year)***
HR (95% CI)			Choueiri 2016[57] (22 May 2015)
Independent radiology review	0.51 (0.41 to 0.62); $p < 0.0001$		
Independent radiology review (discrepancy)	0.52 (0.42 to 0.64); $p < 0.0001$		Choueiri 2016,[57] appendix
Investigator assessed	0.54 (0.44 to 0.65); $p < 0.0001$		
HR (95% CI)			Choueiri 2016[57] (May 2015)
n previous VEGFR-TKIs	independent radiology review		
1	0.52 (0.41 to 0.66)		
≥ 2	0.51 (0.35 to 0.74)		
HR (95% CI)			Choueiri 2016[57] (May 2015)

METEOR			Publication source
Independent radiology review			
MSKCC risk group			
Favourable	0.51 (0.38 to 0.69)		
Intermediate	0.47 (0.35 to 0.65)		
Poor	0.70 (0.42 to 1.16)		
ECOG status			
0	0.46 (0.36 to 0.59)		
1	0.64 (0.46 to 0.90)		
PFS median (95% CI) months			Choueiri 2016[57] (May 2015)
Independent radiology review	7.4 (6.6 to 9.1)	3.9 (3.7 to 5.1)	
Investigator assessed	7.4 (7.3 to 7.8)	5.1 (3.9 to 5.5)	Choueiri 2016,[57] appendix
Number of progression events, *n* (%)			Choueiri 2016[57] (May 2015)
Independent radiology review	180 (55)	214 (66)	
Investigator assessed	196 (59)	233 (71)	
Number of events, *n*/*N* (%) for subgroups based on number of previous VEGFR-TKIs			Choueiri 2016[57] (May 2015)
1	131/235	155/229	
≥ 2	49/95	59/99	
Number of events, *n*/*N* (%) for subgroups based on			Choueiri 2016[57] (May 2015)
MSKCC risk group			
Favourable	79/150	92/150	
Intermediate	74/139	89/135	
Poor	27/54	33/43	
ECOG status			
0	114/226	137/216	
1	66/104	77/112	
OS			
HR (95% CI)	0.67 (0.51 to 0.89); $p = 0.005$		Choueiri 2015[145] (May 2015)
	0.66 (0.53 to 0.83); $p = 0.00026$		Choueiri 2016[57] (December 2015)
	KM figure 2		
HR (95% CI) for subgroups based on number of previous VEGFR-TKIs			Choueiri 2016[57] (December 2015)
1	0.65 (0.50 to 0.85)		
≥ 2	0.73 (0.48 to 1.10)		
HR (95% CI) for subgroups based on			Choueiri 2016[57] (December 2015)
MSKCC risk group			
Favourable	0.66 (0.46 to 0.96)		
Intermediate	0.67 (0.48 to 0.94)		
Poor	0.65 (0.39 to 1.07)		

METEOR			Publication source
ECOG status			
0	0.65 (0.49 to 0.87)		
1	0.72 (0.51 to 1.02)		
OS median (95% CI) months	21.4 (18.7 to not estimable)	16.5 (14.7 to 18.8)	Choueiri 2016[57] (December 2015)
Number of deaths, *n* (%)	140 (42)	180 (55)	Choueiri 2016[57] (December 2015)
Number of deaths, *n*/*N* (%) for subgroups based on number of previous VEGFR-TKIs			Choueiri 2016[57] (December 2015)
1	98/235	130/229	
≥ 2	42/95	50/99	
Number of deaths, *n*/*N* (%) for subgroups based on			Choueiri 2016[57] (December 2015)
MSKCC risk group			
Favourable	48/150	66/150	
Intermediate	64/139	79/135	
Poor	28/41	35/43	
ECOG status			
0	81/226	105/216	
1	59/104	75/112	
Response			
ORR, *n* (%)			Choueiri 2016[57] (May 2015)
Independent radiology review	57 (17) 95% CI 13 to 22	11 (3) 95% CI 2 to 6; $p < 0.0001$	
Investigator assessed	24	4	Choueiri 2016,[57] appendix
Complete response, *n* (%)			Choueiri 2016,[57] appendix
Independent radiology review	0	0	
Investigator assessed	0	0	
PR, *n* (%)			Choueiri 2016,[57] appendix
Independent radiology review	57 (17)	11 (3)	
Investigator assessed	78 (24)	14 (4)	
Stable disease, *n* (%)			Choueiri 2016,[57] appendix
Independent radiology review	216 (65)	203 (62)	
Investigator assessed	209 (63)	205 (63)	
Progressive disease, *n* (%)			Choueiri 2016,[57] appendix
Independent radiology review	41 (12)	88 (27)	
Investigator assessed	29 (9)	87 (27)	
Not evaluable or missing, *n* (%)			Choueiri 2016,[57] appendix
Independent radiology review	16 (5)	26 (8)	
Investigator assessed	14 (4)	22 (7)	

METEOR			Publication source
HRQoL			
Completion Rates (completed/expected at each time point through to week 48 for all instruments)	≥ 75% of patients		Cella 2016[69]
EQ-VAS	–1.32 (17) $n = 317$	–1.27 (16) $n = 304$	NICE TA10075[28]
Mean change from baseline (SD)	MD p-value = 0.921		
EQ-5D-5L index score	–0.02 (0.2) $n = 184$	–0.02 (0.2) $n = 175$	NICE TA10075[28]
Mean change from baseline (SD)	MD p-value = 0.825		
FKSI-19 total	–3.483 (9.8) $n = 319$	–2.21 (9.7) $n = 303$	NICE TA10075[28]
Mean change from baseline (SD)	MD p-value < 0.0001		
Months to deterioration (earlier of death, progression or ≥ 4-point decrease in FKSI-DRS)	5.5	3.7 ($p < 0.0001$ difference)	Cella 2016[69]
Post hoc analysis			
AEs grade ≥ 3,* n *(%)			
n included in safety population	331	322	
Total AEs grade ≥ 3	261 (79)	218 (68)	Choueiri 2016,[57] appendix
Grade 3, 4 and 5 added together, done by the authors in Choueiri 2016[57]			
Total AEs (any grade)	331 (100)	321 (100)	Choueiri 2016[57]
Patients with ≥ 1 SAE grade ≥ 3	130 (39)	129 (40)	Choueiri 2016,[57] abstract
Different criteria to total AEs grade ≥ 3			
Stomatitis	8 (2)	7 (2)	Choueiri 2016[57]
Rash	2 (1)	2 (1)	Choueiri 2016[57]
Fatigue	36 (11)	24 (7)	Choueiri 2016[57]
Asthenia	15 (5)	8 (2)	Choueiri 2016[57]
Diarrhoea	43 (13)	7 (2)	Choueiri 2016[57]
Nausea	15 (5)	1 (< 1)	Choueiri 2016[57]
Vomiting	7 (2)	3 (1)	Choueiri 2016[57]
Cough	1 (< 1)	3 (1)	Choueiri 2016[57]
Dyspnoea	10 (3)	14 (4)	Choueiri 2016[57]
Anaemia	19 (6)	53 (17)	Choueiri 2016[57]
Hypertension	49 (15)	12 (4)	Choueiri 2016[57]
Decreased appetite	10 (3)	3 (1)	Choueiri 2016[57]
PPES	27 (8)	3 (1)	Choueiri 2016[57]
Weight decreased	9 (3)	0	Choueiri 2016[57]
Constipation	1 (< 1)	1 (< 1)	Choueiri 2016[57]
Hypothyroidism	0	1 (< 1)	Choueiri 2016[57]
Dysphonia	2 (1)	0	Choueiri 2016[57]

METEOR			Publication source
Mucosal inflammation	5 (2)	11 (3)	Choueiri 2016[57]
Aspartate aminotransferase concentration increased	5 (2)	1 (< 1)	Choueiri 2016[57]
Pain in extremity	5 (2)	1 (< 1)	Choueiri 2016[57]
Arthralgia	1 (< 1)	4 (1)	Choueiri 2016[57]
Headache	1 (< 1)	1 (< 1)	Choueiri 2016[57]
Dizziness	1 (< 1)	0	Choueiri 2016[57]
Dyspepsia	1 (< 1)	0	Choueiri 2016[57]
Oedema peripheral	0	6 (2)	Choueiri 2016[57]
Hypomagnesaemia	16	0	Choueiri 2016[57]
Proteinuria	8 (2)	2 (1)	Choueiri 2016[57]
Insomnia	0	1 (< 1)	Choueiri 2016[57]
Pyrexia	3 (1)	2 (1)	Choueiri 2016[57]
Pruritus	0	1 (< 1)	Choueiri 2016[57]
Blood creatinine increased	1 (< 1)	0	Choueiri 2016[57]
Hypertriglyceridaemia	4 (1)	10 (3)	Choueiri 2016[57]
Hyperglycaemia	3	16 (5)	Choueiri 2016[57]
Back pain	8 (2)	7 (2)	Choueiri 2016[57]
Abdominal pain	12 (4)	5 (2)	Choueiri 2016[57]
Alanine aminotransferase increased	8 (2)	1 (< 1)	Choueiri 2016[57]

IQR, interquartile range; MET, mesenchymal – epithelial transition; NR, not reported; PD-1, programmed cell death protein 1; PD-L1, programmed death receptor ligand 1; PR, partial response; SAE, serious adverse event; SoD, superoxide dismutase; VEGFR, vascular endothelial growth factor receptor.

Paglino et al.[60]

Paglino *et al.*[60]		Publication source
		Paglino 2013[60]
		Porta 2011[152]
Design		
Study design	Retrospective sequencing study	Paglino 2013[60]
Number of centres and country/countries	Six European centres	Paglino 2013[60]
Cohort recruitment	Data analysed in this study were obtained from the medical records of each individual patient; baseline characteristics, date of start of treatment, dates of progression and time between these treatments were all recorded	Paglino 2013[60]
Recruitment dates	Patients treated between September 2005 and October 2010	Paglino 2013[60]
Length of follow-up	NR	
Source of funding	NR	

Paglino *et al.*[60]			Publication source
Eligibility criteria (inclusion and exclusion)	All patients had received first-line treatment with either sunitinib (50 mg daily, 4 weeks on and 2 weeks off) or sorafenib (400 mg twice daily, continuous dosing), followed by a second-line treatment with a mTORi (either 10 mg everolimus daily continuous dosing or 25 mg temsirolimus intravenous weekly) and, on further progression, with the other multikinase inhibitor (sorafenib or sunitinib), as third-line therapy Patients who were treated with any other agent during the gap between the three drugs treatments were excluded from this analysis		Paglino 2013[60]
Participants and treatment arms	***Intervention: sunitinib***	***Comparator: sorafenib***	
Intervention, method of delivery, dose and frequency	Sequence sorafenib-mTORi-sunitinib Sunitinib (50 mg daily, 4 weeks on and 2 weeks off)	Sequence sunitinib-mTORi-sorafenib Sorafenib (400 mg twice daily, continuous dosing)	Paglino 2013[60]
Concomitant medication(s) or therapies	NR	NR	
Crossover or post-study interventions allowed	NR	NR	
Number of cycles, dose reductions	NR	NR	
Treatment duration	NR	NR	
Number randomised	N/A	N/A	
Number who received study medication	26	14	Paglino 2013[60]
Number withdrawn/discontinued and reasons	NR	NR	
Disease stage and/or metastatic disease	NR	NR	
Previous systemic therapy treatments, *n* (%)			Paglino 2013[60]
First-line TKI	Sunitinib	Sorafenib	
Second-line mTORi	Everolimus/Temsirolimus	Everolimus/Temsirolimus	
Age (years): median (range)	61 (33–75)	63 (44–76)	Paglino 2013[60]
Ethnicity, *n* (%)	NR	NR	
Male, *n* (%)	22 (86)	12 (86)	Paglino 2013[60]
ECOG performance status, *n* (%)			Paglino 2013[60]
0	15 (58)	8 (57)	
1	10 (38)	5 (36)	
2	1 (4)	1 (7)	
MSKCC risk status, *n* (%)			Paglino 2013[60]
Good	12 (46)	8 (57)	
Intermediate	14 (54)	4 (29)	
Poor	0 (0)	2 (14)	
Not reported	0 (0)	0 (0)	

Paglino *et al.*[60]			Publication source
Reported subgroups	OS for age (per year), male vs. female, clear cell vs. non-clear cell, ECOG performance status (1 vs. 0), Motzer's score (1–2 vs. 0), Fuhrman's grade (3–4 vs. 1–2) and Hepatic metastases		Paglino 2013[60]
Reported outcomes			
Primary outcome	Overall PFS, defined as the sum of the PFS achieved under each of the three drugs, excluding any time which may have elapsed between each treatment period. Disease progression during the three treatment periods was determined by radiological assessment using the RECIST approximately every 12 weeks		Paglino 2013[60]
Secondary outcomes	PFS for each line of therapy. PFS for the third line was calculated as the time from the start of the second TKI to the time of disease progression or death. Patients who remained on treatment on the second TKI without disease progression at the end of the study period were censored		Paglino 2013[60]
Results	***Intervention: sunitinib***	***Comparator: sorafenib***	
PFS			
HR (95% CI)	KM, figure 3; $p = 0.2379$		Paglino 2013[60]
HR (95% CI) for subgroups based on prior therapy	NR		
PFS median (IQR) [median (range)], months	3.90 (3.00–13.42)	9.12 (3.50–20.03); $p = 0.2379$	Paglino 2013[60]
Number of progression events, *n* (%)	NR	NR	

IQR, interquartile range; N/A, not applicable; NR, not reported; PS, performance status.

Porta et al.[61]

Porta *et al.*[61]		Publication source
		Porta 2011[61]
		Porta 2010[153]
Design		
Study design	Retrospective sequencing study	Porta 2011[61]
Number of centres and country/countries	12 centres across Italy	Porta 2011[61]
Cohort recruitment	Data analysed in this study were obtained from the medical charts of each individual patient	Porta 2011[61]
Recruitment dates	Patients with RCC who were treated with SoSu or SuSo between March 2004 and April 2009	Porta 2011[61]
Length of follow-up	NR	
Source of funding	Bayer, Berlin, Germany	Porta 2011[61]
Eligibility criteria (inclusion and exclusion)	Patients with RCC who were treated with sorafenib followed by sunitinib or sunitinib followed by sorafenib Patients who were treated with any other agent during the treatment gap between sorafenib and sunitinib therapy were excluded from the study	Porta 2011[61]

Porta *et al.*[61]			Publication source
Participants and treatment arms	***Intervention: sunitinib***	***Comparator: sorafenib***	
Intervention, method of delivery, dose and frequency	50 mg every day on a 4-weeks-on and 2-weeks-off treatment schedule	800 mg/day	Porta 2011[61]
Concomitant medication(s) or therapies	NR	NR	
Crossover or post-study interventions allowed	NR	NR	
Number of cycles, dose reductions	NR	NR	
Treatment duration	NR	NR	
Number randomised	N/A	N/A	
Number who received study medication	90	99	Porta 2011[61]
Number withdrawn/discontinued and reasons	NR	NR	
Disease stage and/or metastatic disease	NR	NR	
Previous systemic therapy treatments, *n* (%)			Porta 2011[61]
TKI	90 (100)	99 (100)	
Systemic therapy prior to TKI			
None	56 (62)	29 (29)	
Cytokine monotherapy	22 (65)	38 (54)	
Cytokines + chemotherapy	9 (26)	27 (39)	
Targeted therapy (bevacizumab + interferon)	1 (3)	2 (3)	
All the above	2 (6)	3 (4)	
Age (years): median (range)	58 (26–78)	60 (32–80)	Porta 2011[61]
Ethnicity, *n* (%)	NR	NR	
Male, *n* (%)	74 (82)	67 (68)	Porta 2011[61]
Performance status, *n* (%)			Porta 2011[61]
ECOG score			
0	64 (71)	71 (72)	
1	24 (27)	25 (25)	
2	2 (2)	3 (3)	
MSKCC			
Good	45 (50)	41 (41)	
Intermediate	35 (39)	26 (26)	
Poor	9 (10)	32 (32)	
Not reported	1 (1)	0 (0)	
Reported subgroups	Subgroups for PFS for second TKI included previous systemic treatment (yes/no), sex (male/female), age (< 65/≥ 65 years), ECOG (0/1), MSKCC (poor/intermediate/good), histology (clear cell/non-clear cell), and centre (Milan/Pavia/Naples/other)		Porta 2011[61]

Porta *et al.*[61]			Publication source
Reported outcomes			
Primary outcome	PFS on the first and second TKI. PFS for the second treatment period was calculated as the time from the start of the second TKI to the time of disease progression or death on the second TKI. The status of disease progression during the first and second TKI was determined by radiological assessment (RECIST) approximately every 12 weeks. Patients who remained on the second TKI without disease progression at the end of the study period were censored from the analysis		Porta 2011[61]
Secondary outcomes			
Outcomes and time points with data reported for subgroups of baseline prognostic scores (e.g. ECOG, MSKCC)	Subgroups for PFS for second TKI ECOG (0/1), MSKCC (poor/intermediate/good)		Porta 2011[61]
Results	***Intervention: sunitinib***	***Comparator: sorafenib***	
PFS			
HR (95% CI)	0.535 (0.387 to 0.740); $p = 0.0002$		Porta 2010[153]
	KM data figure 1c		Porta 2011[61]
HR (95% CI) for subgroups based on prior therapy			
PFS mean ± SD [median (range)], months	7.89 (0.8–26.9)	4.24 (0.1–34.7)	Porta 2010[153]
PFS mean ± SD [median (range)], months			Porta 2011[61]
Systemic therapy prior to first TKI	(8.3)	(4.4) $p = 0.006$	
		KM data figure 1d	
No systemic therapy prior to first TKI	(7.6)	(3.5) $p = 0.054$	
		KM data figure 1d	
HR (95% CI) for subgroups based on prior therapy	0.78 (0.57 to 1.06); $p = 0.115$		Porta 2011[61]
None vs. any systemic treatment prior to first TKI			
HR (95% CI) for subgroups based on prognostic scores			Porta 2011[61]
ECOG 0 vs. 1	0.48 (0.31 to 0.73); $p < 0.001$		
MSKCC good vs. poor	0.44 (0.27 to 0.71); $p < 0.001$		
Good vs. intermediate	0.72 (0.49 to 1.05); $p = 0.084$		
Poor vs. intermediate	0.61 (0.38 to 0.96); $p = 0.033$		

N/A, not applicable; NR, not reported; SoSu, sorafenib sunitinib; SuSo, sunitinib sorafenib.

RECORD-1

RECORD-1		Publication source
		Motzer 2010[93]
		Motzer 2008[64]
		Calvo 2012[74]
		Beaumont 2011[70]
		Hutson 2009[155]
		White 2010[158]
		Figlin 2011[163]
		Kay 2009[157]
		Korhonen 2012[65]
		Wiederkehr 2009[156]
		Escudier 2008[154]
		Figlin 2012[162]
		Osanto 2010[160]
		Porta 2012[161]
		Calvo 2010[159]
		Oudard 2012[164]
Design		
Study design	International, multicentre double-blind, placebo-controlled randomised Phase III trial	Motzer 2008[64]
Number of centres and country/ countries	86 centres in Australia, Canada, Europe, Japan and the USA	Motzer 2008[64]
Cohort recruitment	Patients randomly assigned 2 : 1 to receive either continuous treatments with oral everolimus 10 mg once daily or placebo, both in conjunction with supportive care	Motzer 2008[64]
Recruitment dates	November 2006 to November 2007	Motzer 2008[64] Motzer 2010[93]
Length of follow-up	December 2006 to 15 October 2007 (interim analysis and pre-crossover) December 2006 to February 2008 (final analysis and post crossover); OS data to November 2008	Motzer 2008[64] Motzer 2010[93]
Source of funding	Novartis	Motzer 2008[64]
Eligibility criteria (inclusion and exclusion)	Inclusion: adults (aged ≥ 18 years) with a clear-cell mRCC which had progressed on or within 6 months of stopping treatment with sunitinib or sorafenib, or both drugs. Previous therapy with bevacizumab, interleukin 2 or interferon alfa was also permitted; presence of measurable disease (as per RECIST), a Karnofsky performance status score of ≥ 70%, and adequate bone marrow, hepatic and renal function	Motzer 2008[64]

RECORD-1			Publication source
	Exclusion: previous treatment with mTOR inhibitor therapy (temsirolimus), untreated CNS metastases, or uncontrolled medical conditions (unstable angina, congestive heart failure, recent myocardial infarction or diabetes mellitus)		
Participants and treatment arms	***Intervention: everolimus***	***Comparator: placebo***	
Intervention, method of delivery, dose and frequency	Everolimus, oral, 10 mg once daily (two 5-mg tablets once daily) plus BSC	Placebo once daily plus BSC	Motzer 2008[64]
Concomitant medication(s) or therapies	NR	NR	
Crossover or post-study interventions allowed	NR	Placebo patients could crossover to receive open-label everolimus	Motzer 2008[64] Motzer 2010[93]
		79 patients crossed over, of whom 60 had progressed within 8 weeks of enrolment	
		106/139 (76.2%) patients had crossed over to everolimus before ending the double-blind treatment	
Number of cycles, dose reductions	Each cycle was considered as 28 days of treatment		Motzer 2008[64]
	Doses were delayed or reduced if patients had clinically significant haematological or other AEs that were deemed to be related to everolimus, according to a nomogram described in the protocol and doses were reduced to 5 mg once daily. 92 (34%) patients in the everolimus group and 20 (15%) in the placebo group had a dose interruption, whereas 14 (5%) in the everolimus group and one (< 1%) in the placebo group had a dose reduction with no previous interruption		Motzer 2010[93]
	At least one dose reduction occurred in 7% of everolimus-treated patients and 1% of placebo-treated patients. At least one treatment interruption occurred in 38% of everolimus-treated patients and 11% of placebo-treated patients. Interruptions were because of AEs in 35% and 9%, and laboratory test abnormalities in 3% and 2% of everolimus- and placebo-treated patients, respectively		
Treatment duration, median days (range)	95 (12–315)	57 (21–237)	Motzer 2008[64]
	141 (19–451)	60 (21–195)	Motzer 2010[93]
	Treatment in both groups was continued until disease progression, unacceptable toxicity, death, or discontinuation for any other reason	Treatment in both groups was continued until disease progression, unacceptable toxicity, death, or discontinuation for any other reason	
Number randomised	272	138	Motzer 2008[64]
Extra 5 Japanese patients in Motzer 2010[93] necessary for Japanese Health Authority	277	139	Motzer 2010[93]

RECORD-1			Publication source
Number who received study medication	269	135	Motzer 2008[64]
	274	137	Motzer 2010[93]
Number withdrawn/discontinued and reasons	Total 202	Total 133	Motzer 2010[93]
Disease progression	137	124	
AEs	36	2	
Withdrew consent	13	2	
Death	7	4	
Lost to follow-up	4	0	
Protocol violation	2	1	
Administrative problems	2	0	
Abnormal laboratory values	1	0	
Treatment ongoing at end of double-blind phase	75	6	
Disease stage and/or metastatic disease	Metastatic RCC	Metastatic RCC	
Previous systemic therapy treatments, *n* (%)			
One prior TKI	205/277 (74)	103/139 (74)	Calvo 2012[74]
Two prior TKIs	72/277 (26)	36/139 (26)	Calvo 2012[74]
Sunitinib only	124 (45)	60 (43)	Motzer 2010[93]
Sunitinib as only prior anti-neoplastic therapy	43 (16)	13 (9)	Calvo 2012[74]
Sorafenib only	81 (29)	43 (31)	Calvo 2012[74]
Sunitinib and sorafenib	72 (26)	36 (26)	Motzer 2010[93]
Sunitinib and sorafenib	66 (24)	33 (24)	Figlin 2011[163]
Immunotherapy	179 (65)	93 (67)	Motzer 2010[93]
Chemotherapy	37 (13)	22 (16)	Motzer 2010[93]
Hormone therapy	5 (2)	5 (4)	Motzer 2010[93]
Other systemic therapy	15 (5)	4 (3)	Motzer 2010[93]
Sunitinib or sorafenib	211 (76)	106 (78)	Figlin 2011[163]
	Note results differ between Calvo[74] and Motzer;[93] Motzer data reported when possible	Note results differ between Calvo[74] and Motzer;[93] Motzer data reported when possible	
Age (years): mean ± SD (range)	61 (27–85)	60 (29–79)	Motzer 2008[64]
			Motzer 2010[93]
Ethnicity, *n* (%)	NR	NR	
Male, *n* (%)	216 (78)	106 (76)	Motzer 2010[93]

RECORD-1			Publication source
Performance status (e.g. ECOG, MSKCC, Heng), *n* (%)			Motzer 2010[93]
Karnofsky 100	78 (28)	41 (30)	
Karnofsky 90	98 (36)	53 (38)	
Karnofsky 80	72 (26)	30 (22)	
Karnofsky 70	28 (10)	15 (11)	
Missing	1 (< 1)	0	
MSKCC favourable	81 (29)	39 (28)	
MSKCC intermediate	156 (56)	79 (57)	
MSKCC poor	40 (15)	21 (15)	
Reported subgroups	Interim analysis		Motzer 2008[64]
	• PFS by MSKCC risk, and previous systemic treatment (stratification factors)		Motzer 2010[93]
	Final analysis		
	• OS by MSKCC		
Reported outcomes			
Primary outcome	PFS (defined as time from randomisation to the first documentation of disease progression or death, from any cause) assessed via blinded independent central review		Motzer 2008[64]
Secondary outcomes	OS, objective tumour response rate, disease-related symptoms, quality of life, AEs		Motzer 2008[64]
Outcomes and time points with data reported for subgroups of prior baseline therapies	PFS by previous systemic treatment (interim analysis at October 2007 – pre-crossover)		Motzer 2008[64]
	PFS by number of prior treatments		Calvo 2012[74]
Outcomes and time points with data reported for subgroups of baseline prognostic scores (e.g. ECOG, MSKCC)	PFS by previous systemic treatment (interim analysis at October 2007 – pre-crossover)		Motzer 2008[64]
	PFS by prognostic score MSKCC		Calvo 2012[74]
	OS by MSKCC (final analysis at February 2008 – post crossover)		Motzer 2010[93]
Results	***Intervention: everolimus*** **n = *272*** **n = *277***	***Comparator: placebo*** **n = *138*** **n = *139***	***Publication and data cut-off point (month, year)*** ***Motzer 2008[64]/2010[93]***
PFS			
HR (95% CI)	0.30 (0.22 to 0.40); $p < 0.001$		Motzer 2008[64] (October 2007)
	0.33 (0.25 to 0.43); $p < 0.001$ (independent central review)		Motzer 2010[93] (February 2008)
HR (95% CI) for subgroups based on prior therapy			Motzer 2008[64] (October 2007)
Sunitinib prior treatment:	0.30; $p < 0.0001$ ($n = 184$)		Motzer 2008[64] (October 2007)
Sorafenib prior treatment	0.29; $p < 0.0001$ ($n = 119$)		

RECORD-1			Publication source
Sunitinib and sorafenib prior treatment:	0.32 (0.19 to 0.54) ($n = 108$)		Motzer 2010[93] (February 2008)
Sunitinib prior treatment	0.34 (0.23 to 0.51); $p < 0.001$		Calvo 2012[74]
Sorafenib prior treatment	0.25 (0.16 to 0.42); $p < 0.001$		Calvo 2012[74]
Sunitinib as the only prior anti-neoplastic therapy	0.22 (0.09 to 0.55); $p < 0.001$		Calvo 2012[74]
Sorafenib as the only prior anti-neoplastic therapy	0.35 (0.14 to 0.88); $p < 0.001$		Calvo 2012[74]
Prior bevacizumab	0.3 (0.13 to 0.68); $p = 0.001$		Hutson 2009[155]
HR (95% CI) for subgroups based on number of prior therapy			
1 previous TKI therapy:	0.32 (0.24 to 0.43); $p < 0.001$		Calvo 2012[74]
	0.31 (0.23 to 0.42); $p < 0.001$		Figlin 2011[163]
2 previous TKI therapy	0.32 (0.19 to 0.54); $p < 0.001$		Calvo 2012[74]
	0.37 (0.22 to 0.63); $p < 0.001$		Figlin 2011[163]
	(Note: different numbers reported in Calvo 2012 and Figlin 2011 for same outcomes; Figlin is abstract only)		
HR (95% CI) for subgroups based on MSKCC risk			Motzer 2008[64] (October 2007)
MSKCC favourable:	0.35; $p < 0.0001$ ($n = 118$)		Calvo 2012[74]
MSKCC intermediate:	0.25; $p < 0.0001$ ($n = 231$)		
MSKCC poor:	0.39; $p = 0.009$ ($n = 61$)		
MSKCC favourable:	0.31 (0.19 to 0.50); $p < 0.001$ ($n = 120$)		
MSKCC intermediate:	0.32 (0.22 to 0.44); $p < 0.001$ ($n = 235$)		
MSKCC poor:	0.44 (0.22 to 0.85); $p = 0.007$ ($n = 61$)		
PFS, median (95% CI), months	4.0 (3.7 to 5.5)	1.9 (1.8 to 1.9)	Motzer 2008[64] (October 2007)
	4.9 (4.0 to 5.5)	1.9 (1.8 to 1.9)	Motzer 2010[93] (February 2008)
PFS median (range), months for subgroups based on prior therapy			Calvo 2012[74]
Sunitinib prior treatment	3.9	1.8	Calvo 2012[74]
Sorafenib prior treatment:	5.9	2.8	Motzer 2010[93] (February 2008)
Sunitinib and sorafenib prior treatment:	4.0	1.0	Calvo 2012[74]
Sunitinib as the only prior anti-neoplastic therapy	4.6	1.8	Calvo 2012[74]
Sorafenib as the only prior anti-neoplastic therapy	3.8	1.9	Hutson 2009[155]
Prior bevacizumab	5.75 (3.52 to 6.90)	1.77 (1.02 to 3.78)	Calvo 2012[74]
One previous TKI therapy:	5.42 (95% CI 4.30 to 5.82)	1.87 (95% CI 1.84 to 2.14)	Figlin 2011[163]
Two previous TKI therapy	3.78 (95% CI 3.25 to 5.13) [4.0 in Calvo 2012[74]]	1.87 (95% CI 1.77 to 3.06) [1.8 in Calvo 2012[74]]	Calvo 2012[74]

RECORD-1			Publication source
PFS median (range), months for subgroups based on MSKCC risk			Calvo 2012[74]
MSKCC favourable	5.8 (n = 81)	1.9 (*n* = 39)	
MSKCC intermediate	4.5 (*n* = 156)	1.8 (*n* = 79)	
MSKCC poor	3.6 (*n* = 40)	1.8 (*n* = 21)	
Number of progression events, *n* (%)	101/372 (37)	90/138 (65)	Motzer 2008[64] (October 2007)
Progression	85	82	
Death	16	8	
Censored	171	48	
Overall survival			
HR (95% CI)	0.83 (0.50 to 1.37); *p* = 0.23)		Motzer 2008[64] (October 2007)
	0.87 (0.65 to 1.15); *p* = 0.162		Motzer 2010[93] (November 2008)
	Note: before November 2008, 111 (80%) of 139 placebo patients received open-label everolimus. Survival results likely to be confounded by crossover to everolimus		
Crossover corrected (IPCW analysis)	0.55 (0.31 to 0.97)		Wiederkehr 2009[156] (February 2008)
Crossover corrected (RPSFTM)	0.60 (0.22 to 1.65)		Korhonen 2012[65] (November 2008)
OS, median (range), months	Not reached	8.8	Motzer 2008[64]
	14.78	14.39	Motzer 2010[93]
Crossover corrected (IPCW analysis)	NR	NR	Wiederkehr 2009[156]
Crossover corrected (RPSFTM)	14.4	10.0	Korhonen 2012[65] (November 2008)
HR (95% CI) for subgroups based on prior therapy	NR		
Number of deaths, *n* (%)	42/272 (15)	26/138 (19)	Motzer 2008[64] (October 2007)
	146 (52.7)	75 (54.0)	Korhonen 2012[65] (November 2008)
Number of deaths, *n* (%) for subgroups based on prior therapy	NR	NR	
Response			
ORR, *n* (%)	NR	NR	
Complete response rate, *n* (%)	0	0	Motzer 2010[93] (February 2008)
PR rate, *n* (%)	3 (1) (*n* = 272)	0 (*n* = 138)	Motzer 2008[64] (October 2007)
	5 (1.8)	0	Motzer 2010[93] (February 2008)

RECORD-1					Publication source
Stable disease, *n* (%)	171 (63) (*n* = 272)		44 (32) (*n* = 138)		Motzer 2008[64] (October 2007)
	185/277 (66.8)		45/139 (32.4)		Motzer 2010[93] (February 2008)
Time to response, months mean ± SD [median (range)]	NR		NR		
Duration of response (months)	NR		NR		
Progressive disease, *n* (%)	53 (19) (*n* = 272)		63 (46) (*n* = 138)		Motzer 2008[64] (October 2007)
Disease could not be assessed	45 (17) (*n* = 272)		48 (35) (*n* = 138)		Motzer 2008[64] (October 2007)
HRQoL					
Completion rates, %					Beaumont 2011[70]
Baseline	87		92		
3 months	60				
6 months	32				
EORTC QLQ-C30					
HR (95% CI)	0.94 (0.64 to 1.39)				Motzer 2008[64]
Physical functioning scale	0.84 (0.75 to 0.94); *p* = 0.001				Beaumont 2011[70]
EORTC QLQ-C30					
HR (95% CI)	1.02 (95%CI: 0.70 to 1.50)				Motzer 2008[64]
Global health status/quality of life score	0.85 (0.76 to 0.96); *p* = 0.006				Beaumont 2011[70]
FKSI-DRS					Beaumont 2011[70]
HR (95% CI)	0.82 (0.74 to 0.92); *p* = 0.001				
Time to definitive deterioration of the FKSI-DRS risk score by 2 U					Motzer 2010[93]
HR (95%)	0.75 (0.53 to 1.06); *p* = 0.053				
Time to definitive deterioration of the FKSI-DRS risk score by 2 units, median (months)	4.76 (*n* = 277)		3.84 (*n* = 139)		Motzer 2010[93]
AEs grade ≥ 3, *n* (%)					
	Everolimus (*n* = 274)		Placebo (*n* = 137)		
	Grade 3	Grade 4	Grade 3	Grade 4	
Stomatitis (including apthous stomatitis, mouth ulceration and tongue ulceration)	4	< 1	0	0	Motzer 2010[93]
Infections (all infections including pneumonia, aspergillosis, candidiasis and sepsis)	7	3	1	0	Motzer 2010[93]
Asthenia	3	< 1	4	0	Motzer 2010[93]
Fatigue	5	0	3	< 1	Motzer 2010[93]

RECORD-1					Publication source
Diarrhoea	1	0	0	0	Motzer 2010[93]
Cough	< 1	0	0	0	Motzer 2010[93]
Rash	1	0	0	0	Motzer 2010[93]
Nausea	1	0	0	0	Motzer 2010[93]
Anorexia	1	0	< 1	0	Motzer 2010[93]
Peripheral oedema	< 1	0	< 1	0	Motzer 2010[93]
Dyspnoea	6	1	3	0	Motzer 2010[93]
Vomiting	2	0	0	0	Motzer 2010[93]
Pyrexia	< 1	0	0	0	Motzer 2010[93]
Mucosal inflammation	1	0	0	0	Motzer 2010[93]
Headache	< 1	< 1	< 1	0	Motzer 2010[93]
Epistaxis	0	0	0	0	Motzer 2010[93]
Pruritus	< 1	0	0	0	Motzer 2010[93]
Pneumonitis (includes interstitial lung disease, lung infilitration, pneumonitis, pulmonary alveolar haemorrhage, alveolitis and pulmonary toxicity)	4	0	0	0	Motzer 2010[93]
Dry skin	< 1	0	0	0	Motzer 2010[93]
Dysgeusia	0	0	0	0	Motzer 2010[93]
Pain in extremity	1	0	0	0	Motzer 2010[93]
Haemoglobin levels decreased	12	1	5	< 1	Motzer 2010[93]
Lymphocyte levels decreased	16	2	5	0	Motzer 2010[93]
Platelet levels decreased	1	0	0	< 1	Motzer 2010[93]
Neutrophil levels decreased	0	< 1	0	0	Motzer 2010[93]
Cholesterol levels increased	4	0	0	0	Motzer 2010[93]
Triglyceride levels increased	< 1	0	0	0	Motzer 2010[93]
Glucose levels increased	15	< 1	1	0	Motzer 2010[93]
Creatinine levels increased	1	0	0	0	Motzer 2010[93]
Phosphate levels decreased	6	0	0	0	Motzer 2010[93]
Aspartate transaminase levels increased	< 1	< 1	0	0	Motzer 2010[93]
Alanine transaminase levels increased	1	0	0	0	Motzer 2010[93]
Bilirubin levels increased	< 1	< 1	0	0	Motzer 2010[93]
Clinical suspicion of pneumonitis	10 (3.6)	0	0	0	White 2010[112]

NR, not reported; PR, partial response.

SWITCH

SWITCH			
			Eichelberg 2015[56] Michel 2012[165] Eichelberg 2012[166]
Design			
Study design	Randomised, open-label, Phase III trial with crossover study phase (sequential sorafenib/sunitinib and sunitinib/sorafenib trial)		Eichelberg 2015[56]
Number of centres and country/ countries	72 centres in Germany, Austria and the Netherlands		Eichelberg 2015[56]
Recruitment dates	February 2009 to December 2011		Eichelberg 2015[56]
Length of follow-up	February 2009 to 15 August 2013 with updated post hoc OS analysis in January 2014. Mean follow-up 10.3 months (at August 2013 data cut-off point; last treatment to end of follow-up)		Eichelberg 2015[56]
Source of funding	German Cancer Society and grant from Bayer		Eichelberg 2015[56]
Eligibility criteria (inclusion and exclusion)	Inclusion: adults aged 18–85 years, amRCC; unsuitable for cytokine therapy; no prior systemic therapy; ECOG performance status 0 or 1; one or more measurable lesions by CT or MRI according to RECIST; favourable or intermediate MSKCC risk score; and adequate bone marrow, liver, and renal function Exclusion: unstable or severe cardiac disease; active, clinically serious infections; and symptomatic metastatic brain tumours		Eichelberg 2015[56]
Participants and treatment arms	***Intervention: sunitinib***	***Comparator: sorafenib***	
Intervention, method of delivery, dose and frequency	Sorafenib 400 mg twice daily followed by sunitinib 50 mg once daily (4 weeks on, 2 weeks off) First-line treatment continued until disease progression according to RECIST or intolerable toxicity (after unsuccessful dose reduction/interruption). There was a treatment-free crossover period of 1–4 weeks after first-line treatment to avoid additive toxicity. Patients who discontinued first-line treatment because of toxicity began second-line treatment only after non-haematological toxicity had resolved to grade 1 and haematological toxicity to grade 2. Patients who refused further first-line treatment because of toxicity could begin second-line treatment	Sunitinib 50 mg once daily followed by sorafenib 400 mg twice daily (4 weeks on, 2 weeks off) First-line treatment continued until disease progression according to RECIST or intolerable toxicity (after unsuccessful dose reduction/interruption). There was a treatment-free crossover period of 1–4 weeks after first-line treatment to avoid additive toxicity. Patients who discontinued first-line treatment because of toxicity began second-line treatment only after non-haematological toxicity had resolved to grade 1 and haematological toxicity to grade 2. Patients who refused further first-line treatment because of toxicity could begin second-line treatment	Eichelberg 2015[56]

SWITCH			
Concomitant medication(s) or therapies	NR	NR	
Crossover or post-study interventions allowed	Patients crossed over from sorafenib in first line to sunitinib in second line: *n* = 103 (57%)	Patients crossed over from sunitinib in first line to sorafenib in second line: *n* = 76 (42%)	Eichelberg 2015[56]
Number of cycles, dose reductions	Dose changes were allowed in order to manage AEs. The sorafenib dose could be reduced to 400 mg once daily and then 400 mg every other day. The sunitinib dose could be reduced to 37.5 mg once daily and then 25 mg once daily		Eichelberg 2015[56]
Number of cycles	NR	NR	Eichelberg 2015[56]
Dose reductions: second-line sunitinib or sorafenib, *n* (%)	24 (23)	35 (46)	
Treatment duration, mean (± SD), weeks	28.2 (29.6)	16.0 (15.2)	Eichelberg 2015[56]
Number randomised	N/A	N/A	
Number who received study medication	103	76	Eichelberg 2015[56]
Number withdrawn/discontinued and reasons			Eichelberg 2015[56]
Total discontinued	91	71	
Progressive disease	67	53	
AEs	7	6	
Death	3	3	
Withdrew	NR	4	
Health deterioration	5	1	
Other reasons	9	4	
Disease stage and/or metastatic disease	Metastatic RCC	Metastatic RCC	
Previous systemic therapy treatments, *n* (%)			Eichelberg 2015[56]
First line			
Interferon-alpha	2 (1.1)	5 (2.7)	
IL-2	1 (0.5)	3 (1.6)	
Second line			
Sorafenib	103 (57)	N/A	
Sunitinib	N/A	76 (42)	
Age (years): median (range)	62 (40–81)	63 (41–81)	Eichelberg 2015[56]
Ethnicity, *n* (%)	NR	NR	
Male, *n* (%)	81 (79)	56 (74)	Eichelberg 2015[56]
Performance status, *n* (%)			Eichelberg 2015[56]
ECOG score			
0	75 (73%)	48 (63%)	
1	27 (26%)	28 (37%)	
2	0	0	
Missing	1 (1%)	0	

SWITCH			
MSKCC			
Favourable	47 (46%)	38 (50%)	
Intermediate	56 (54%)	38 (50%)	
High	0	0	
Missing	0	0	
Reported subgroups	Combined first- and second-line OS and PFS by ECOG performance status and MSKCC score. In addition, age and gender. All subgroups were post hoc		
Reported outcomes			Eichelberg 2015[56]
Primary outcome	PFS (time from randomisation to confirmed progression or death during second-line therapy, i.e. combined first and second line)		
Secondary outcomes	Combined first- and second-line OS (time from randomisation to time of death from any cause); first-line PFS, second-line PFS, best objective response (CR, PR, stable disease), ORR, disease control rate, AEs, total time to progression (from randomisation to confirmed progression during second-line), time to first treatment failure (randomisation to progression, death or discontinuation)		
Outcomes and time points with data reported for subgroups of prior baseline therapies	NR		
Outcomes and time points with data reported for subgroups of baseline prognostic scores (e.g. ECOG, MSKCC)	Combined first- and second-line OS and PFS by ECOG performance status and MSKCC score at data cut-off point on 15 August 2013		
Results	***Intervention: sunitinib***	***Intervention: sorafenib***	
PFS			
HR (95% CI)	0.55 (0.41 to 0.74); $p < 0.001$ (KM plot figure 3)		Eichelberg 2015[56]
HR (95% CI) for subgroups based on prior therapy	N/A		
PFS median (range), months	5.4 (3.0–5.5)	2.8 (2.7–2.9)	Eichelberg 2015[56]
PFS mean ± SD [median (range)], months for subgroups based on prior therapy	N/A	N/A	
Number of progression events, *n* (%)	NR	NR	
Response: data for second line			
ORR, *n* (%)	18 (17)	5 (6.6)	Eichelberg 2015[56]
CR rate, *n* (%)	1 (1)	1 (1.3)	Eichelberg 2015[56]
PR rate, *n* (%)	17 (17)	4 (5.3)	Eichelberg 2015[56]
Stable disease, *n* (%)	32 (31)	19 (25)	Eichelberg 2015[56]
Time to response, months	NR	NR	
Mean ± SD			
Median (range)			

SWITCH			
Duration of response, months	NR	NR	
Mean ± SD			
Median (range)			
Disease control rate, *n* (%)	50 (49)	24 (32)	Eichelberg 2015[56]
AE grade ≥ 3, *n* (%) (data are second line only)			Eichelberg 2015[56]
Total AEs grade ≥ 3	53 (51) (*n* = 103)	27 (36) (*n* = 76)	
Total TEAEs (any grade)	90 (87)	64 (84)	
AEs leading to withdrawal	20 (19)	15 (20)	
Any serious AEs	43 (42)	19 (25)	
Death-related AEs	1 (1.0)	2 (2.6)	
Stomatitis	0	0	
Rash	0	1 (1.3)	
Fatigue	3 (2.9)	0	
Alopecia	NR	NR	
Diarrhoea	2 (1.9)	3 (3.9)	
Hand-foot skin reaction	5 (4.9)	5 (6.6)	
Nausea	1 (1)	1 (1.3)	
Pain	4 (3.9)	0	
Hypertension	3 (2.9)	2 (2.6)	
Loss of appetite	2 (1.9)	0	
Thrombocytopenia	0	0	

CR, complete response; N/A, not available; IL-2, interleukin 2; NR, not reported; PR, partial response; PS, performance status.

Vogelzang et al.[62]

Vogelzang *et al.*[62]		
		Vogelzang 2016[62]
		Pal 2016[167]
Design		
Study design	Retrospective observational study	Vogelzang 2016[62]
Number of centres and country/ countries	USA	Vogelzang 2016[62]
Cohort recruitment	Medical oncologists and haematologists/oncologists from a nationwide panel in the USA who had treated three or more mRCC patients in the past year. After screening, physicians randomly selected and abstracted data for up to five patient charts that met the prespecified inclusion criteria A standardised electronic case report form was used to extract relevant chart information	Vogelzang 2016[62] Pal 2016[167]

Vogelzang *et al.*[62]			
Recruitment dates	Medical oncologists and haematologists/oncologists were screened from June 2014 to July 2014 Included patient initiated treatment with everolimus or axitinib between 1 February 2012 and 31 January 2013		Vogelzang 2016[62] Pal 2016[167]
Length of follow-up	Mean (SD) follow-up time was 15 (7) months for patients on everolimus and 13 (7) months for patients on axitinib		Vogelzang 2016[62]
Source of funding	Novartis		Vogelzang 2016[62]
Eligibility criteria (inclusion and exclusion)	Patients were required to be at least 18 years of age, to have had a mRCC diagnosis, to have received a TKI (sunitinib, sorafenib or pazopanib) as first targeted therapy and to have discontinued that therapy for medical reasons (e.g. drug intolerance, disease progression, non-response without progression). In addition, patients were required to have subsequently initiated either everolimus or axitinib as second targeted therapy between 1 February 2012 and 1 January 2013 Patients were excluded if they initiated second targeted therapy as part of a clinical research protocol, used IL-2 prior to or in combination with the first TKI for the treatment of mRCC, or used combination therapy with two or more targeted agents as first or second targeted therapy		Vogelzang 2016[62]
Participants and treatment arms	***Intervention: everolimus***	***Comparator: axitinib***	
Intervention, method of delivery, dose and frequency	10 mg once daily	5 mg twice daily	Pal 2016[167]
Concomitant medication(s) or therapies	NR	NR	
Crossover or post-study interventions allowed (including number of patients)	N/A	N/A	
Number of cycles, dose reductions, *n* (%)			Pal 2016[167]
Recommended dose	297 (91)	106 (84)	
Higher dose	7 (2)	18 (14)	
Lower dose	22(7)	3 (2)	
No change in dose	300 (92)	115 (87)	
Dose escalation	3 (1)	14 (11)	
Dose de-escalation	22 (7)	3 (2)	
Treatment duration	NR	NR	
Number randomised	N/A	N/A	
Number who received study medication	325	127	Vogelzang 2016[62]
Number withdrawn/discontinued and reasons	NR	NR	
Disease stage and/or metastatic disease			Vogelzang 2016[62]
Metastasised RCC at initial diagnosis, *n* (%)	165 (51)	75 (59)	
Clear-cell RCC, *n* (%)	274 (84)	111 (87)	

Vogelzang *et al.*[62]			
Previous systemic therapy treatments, *n* (%)			Vogelzang 2016[62]
Sunitinib	239 (74)	89 (70)	
Pazopanib	59 (18)	33 (26)	
Sorafenib	27 (8)	5 (4)	
Age (years): mean (SD)	61 (9)	60 (9)	Vogelzang 2016[62]
Ethnicity, *n* (%)	NR	NR	
Male, *n* (%)	229 (70)	82 (65)	Vogelzang 2016[62]
Performance status			Vogelzang 2016[62]
ECOG, *n* (%)			
0	99 (30)	43 (34)	
1	163 (50)	64 (50)	
≥ 2	63 (19)	20 (16)	
Reported subgroups	NR		
Reported outcomes			
Outcomes	OS was defined as the time from the initiation of second targeted therapy to death from any cause. Patients without a recorded date of death at the time of medical records review were censored at the last recorded follow-up date PFS was defined as the time from the initiation of second targeted therapy to progression or death, whichever came first. Patients without a recorded date of progression or death were censored at the last recorded follow-up date Progression was determined by physicians based on radiographic evidence indicating progression of tumour lesions or occurrence of new lesions, physical exams, worsening performance status, worsening hypercalcaemia, growth of subcutaneous mass or palpable mass, or cancer-related symptoms (e.g. increased pain, fever and weight loss)		Vogelzang 2016[62] Pal 2016[167]
Outcomes for subgroups of baseline prognostic scores	PFS and OS presented with subgroups based on performance status, ECOG score		Vogelzang 2016[62]
Results	***Intervention: everolimus***	***Comparator: axitinib***	
PFS			
HR (95% CI)			Vogelzang 2016[62]
Unadjusted	1.07 (0.70 to 1.64); $p = 0.742$		
Adjusted[a]	1.16 (0.85 to 1.59); $p = 0.352$		
HR (95% CI) for subgroups based on			Vogelzang 2016[62]
ECOG score			
1	1.61 (1.15 to 2.24); $p = 0.005$		
≥ 2	2.53 (1.67 to 3.83); $p < 0.001$		
PFS mean ± SD [median (range)] months	NR	NR	
PFS mean ± SD [median (range)] months, for subgroups based on performance status	NR	NR	

Vogelzang *et al.*[62]			
Number of progression events *n* (%)	174 (54)	59 (46)	Vogelzang 2016[62]
Number of progression events, *n* (%) at 12 months	60	56	Pal 2016[167]
Overall survival			
HR (95% CI)			Vogelzang 2016[62]
Unadjusted	1.02 (0.67 to 1.55); $p = 0.938$		
Adjusted	1.16 (0.74 to 1.82); $p = 0.531$		
HR (95% CI) for subgroups based			Vogelzang 2016[62]
ECOG score			
1	1.79 (1.08 to 2.97); $p = 0.025$		
≥ 2	3.73 (2.02 to 6.87); $p = 0.001$		
Number of deaths, *n* (%)	83 (26)	29 (23)	Vogelzang 2016[62]
Number of deaths, *n* (%) for subgroups based on prior therapy	NR	NR	

IL-2, interleukin 2; N/A, not applicable; NR, not reported.
a Multivariable Cox PHs models adjusted for age, sex, metastasized RCC at initial diagnosis, prior nephrectomy, type and duration of first targeted therapy, clinical benefit while on first targeted therapy (physician assessed yes/no), occurrence of progression while on first targeted therapy (physician assessed yes/no), duration of mRCC at second targeted therapy initiation, sites of metastases, clear-cell RCC histology, comorbid hypercholesterolemia, ECOG performance status, and years of practice of the treating physician.

Wong et al.[63]

Wong *et al.*[63]		
		Wong 2014[63]
		Wong 2014,[63] appendix
Design		
Study design	Retrospective observational study	Wong 2014[63]
Number of centres and country/countries	Number of centres not reported; country: USA	Wong 2014[63]
Cohort recruitment	Physicians were recruited from a panel covering 12% of specialists in oncology or oncology/haematology in the USA. Among 1575 physicians invited during 2009/11 to 2011/12, 159 agreed to participate and had eligible patient charts. These physicians were instructed to identify all charts meeting the selection criteria to randomly select up to five of those charts for data extraction. The study sponsor and the authors did not have any influence in the process of sampling physicians and the participating physicians were blinded to the identity of the study sponsor	Wong 2014[63]
Recruitment dates	Physicians were recruited between 2009/11 and 2011/12	Wong 2014[63]
Length of follow-up	12.9 months for everolimus and 12.1 months for sorafenib	Wong 2014[63]
Source of funding	Partially funded by Novartis	Wong 2014[63]

Wong *et al.*[63]			

Eligibility criteria (inclusion and exclusion)	Inclusion criteria
	
	• Adult patients (aged ≥ 18 years) with a diagnosis of mRCC
	• The patient's first targeted therapy for mRCC was a TKI (sunitinib, sorafenib or pazopanib)
	• The patient discontinued the first-line TKI for medical reasons (e.g. progression, no response, tolerability) and subsequently initiated one of the following as a second targeted therapy:
	  ○ A different TKI (i.e. different from the first-line treatment, other than axitinib, which was excluded) or
	  ○ An mTOR inhibitor (everolimus)
	• The date of initiation of the second targeted therapy was between 10/2009 and 06/2010. This time window was chosen to allow 6 months after FDA approval of all study drugs and sufficient minimal follow-up time for assessing OS and PFS
	• The patient's medical records are available for review from the initiation of the first-line TKI therapy until the most recent follow-up or death
	
	Exclusion criteria
	
	• Participation in any clinical trials of investigational treatments for mRCC prior to or on initiation of the second targeted therapy
	• Use of an mTOR inhibitor or bevacizumab prior to first-line TKI for the treatment of mRCC
	• Use of combination therapy with two targeted agents prior to initiating the second targeted therapy
	• Use of high dose IL-2 (i.e. on average 600,000 U/kg per day) prior to the first TKI therapy
	• Initiation of the second targeted therapy in combination with another targeted therapy
			Wong 2014[63]

Participants and treatment arms	***Intervention: everolimus***	***Comparator: sorafenib***	
Intervention, method of delivery, dose and frequency	NR	NR	
Concomitant medication(s) or therapies	NR	NR	
Post-study interventions allowed	NR	NR	Wong 2014[63]
Initiated third targeted therapy among patients discontinuing second targeted therapy, *n* (%)			
TKI	33 (84.6)	4 (26.7)	
mTOR	5 (12.8)	11 (73.3)	
Number of cycles			Wong 2014[63]
Rate of dose adjustment	21 (9.0)	29 (23.6)	
Treatment duration	NR	NR	
Number randomised	N/A	N/A	
Number who received study medication	233	123	Wong 2014[63]

Wong *et al.*[63]			
Number withdrawn/discontinued and reasons	NR	NR	
Discontinued the second targeted therapy, *n* (%)	108 (48.2)	60 (50.0)	Wong 2014[63]
Number of metastatic sites, median (range)	2 (1–6)	2 (1–4)	Wong 2014[63]
Previous systemic therapy[a] treatments, *n* (%)	1 (0.4)	2 (1.6)	Wong 2014[63]
First targeted therapy, *n* (%)			Wong 2014[63]
Sunitinib	186 (79.8)	122 (99.2)	
Sorafenib	32 (13.7)[a]	0 (0.0)	
Pazopanib	15 (6.4)	1 (0.8)	
Age (years): median (range)	64 (36–82)	66 (34–83)	Wong 2014[63]
Ethnicity (white), *n* (%)	191 (82.0)	97 (78.9)	Wong 2014[63]
Male, *n* (%)	164 (70.4)	88 (71.5)	Wong 2014[63]
Performance status, KPS[b] *n* (%)			Wong 2014[63]
70–100%	184 (80.7)	98 (83.8)	
0–60%	44 (19.3)	19 (16.2)	
Reported subgroups	Duration of first-line TKI		
Reported outcomes			
Primary outcome	OS was defined as the time from initiation of the second targeted therapy to death from any cause		Wong 2014[63]
Secondary outcomes	PFS was defined as the time from initiation of the second targeted therapy to disease progression or death. Progression was assessed by the treating physician, as recorded in the medical records, based on radiographic evidence, physical examinations or worsening cancer-related symptoms		Wong 2014[63]
Outcomes and time points with data reported for subgroups of prior baseline therapies	Adjusted comparison of OS between second-line everolimus and sorafenib, subgroup analysis stratified by duration of first-line TKI duration (≥ 6 months and < 6 months)		Wong 2014[63]
Outcomes and time points with data reported for subgroups of baseline prognostic scores (e.g. ECOG, MSKCC)	NR		
Results	***Intervention: everolimus***	***Comparator: sorafenib***	
PFS			
Unadjusted HR (95% CI)	1.01 (0.75 to 1.37); $p = 0.931$		
Adjusted[c] HR (95% CI)	0.76 (0.55 to 1.04); $p = 0.090$		Wong 2014[63]
	0.75[d] (0.53 to 1.07); $p = 0.110$		Wong 2014,[63] appendix
HR (95% CI) for subgroups based on prior therapy	NR		
PFS mean ± SD [median (range)] months	10.1	8.6	Wong 2014[63]
PFS mean ± SD [median (range)] months, for subgroups based on prior therapy	NR	NR	

Wong *et al.*[63]			
Number of progression events, *n* (%)	138 (59)	70 (57)	Wong 2014[63]
Overall survival			
Adjusted[c] HR, (95% CI)	0.66 (0.44 to 0.99); $p = 0.045$		Wong 2014[63]
	0.65[d] (0.42 to 0.99); $p = 0.047$		Wong 2014,[63] appendix
Unadjusted HR (95% CI)	1.05 (0.72 to 1.51); $p = 0.809$		Wong 2014[63]
HR (95% CI) for subgroups based on prior therapy	NR		
Number of deaths, *n* (%)	100 (42.9)	48 (39.0)	Wong 2014[63]
Adjusted[c] median OS months	19.0	13.8	Wong 2014[63]
Number of deaths, *n* (%), for subgroups based on prior therapy	NR	NR	

FDA, Food and Drug Administration; IL-2, interleukin 2; KPS, Karnofsky performance status; N/A, not applicable; NR, not reported.

a Systemic therapy prior to initiation of the first targeted therapy includes interferon, IL-2 and chemotherapy.

b KPS score of 100% indicates a perfect performance status and 0% indicates death.

c Multivariable Cox PHs models adjusted for age, gender, race, whether metastasis was present at initial diagnosis, duration of mRCC, type of first targeted therapy, response to and duration of first targeted therapy, treatments received before first targeted therapy, comorbidities, number and sites of metastasis, sarcomatoid differentiation, non-clear-cell RCC, KPS, physician's practice setting and year of practice. Patients with missing baseline values were excluded from the analyses.

d Difference in analysis and results compared with the full publication is unclear.

Economic evaluation

Identified economic evaluations in people with renal cell carcinoma patients

Author, year, country	Perspective, discounting and cost year	Model type, cycle length, time horizon	Patient population	Intervention/ comparator	Costs included	Outcomes (source)	Results
Paz-Ares *et al.*, 2010, Spain[83]	Spanish NHS, 3.5%, 2007	Markov model with 4-week cycles, 10 years	Patients with mRCC who did not respond to, were intolerant to or experienced disease progression on, IL-2 or IFN-alpha	Sunitinib (50 mg, daily) for 4 weeks followed by a 2-week rest period compared with BSC	Analgesics, medical visits, monitoring, TC, disease progression costs and AE costs	OS and PFS (open-label single-arm Phase II sunitinib trial)	€34,196 per QALY
Petrou *et al.*, 2014, Cyprus[77]	Cypriot health-care payer, 3.5%, 2012	Markov model with monthly cycles, 10 years	Patients with RCC. Median age of patients was 58 years. First-line treatment was cytokine therapy in 85% of patients	Sorafenib second line	Medical and pharmaceutical costs	OS and PFS (Phase III randomised clinical trial)	€102,616 per QALY
Purmonen *et al.*, 2008, Finland[84]	Finnish health-care payer, 5%, 2005	Markov model with monthly cycles, 5 years	Patients with mRCC who have experienced failure on prior cytokine-based therapy	Sunitinib (50 mg, daily) for 4 weeks followed by a 2-week rest period compared with BSC including palliative biochemotherapy	Medications, examinations, hospital unit costs	OS and PFS (two single-arm Phase II sunitinib trials)	€43,698 per QALY
Petrou, 2014, Cyprus[76]	Cypriot health-care payer, 3.5%, cost year not reported	Markov model with monthly cycles, 10 years	Patients with mRCC who have failed on first-line therapy. Median age of patients was 61 years and 54% were male	Axitinib second line (5 mg, twice daily) compared with sorafenib (400 mg twice daily)	Pharmaceutical and medical costs	OS and PFS (AXIS trial)	€87,936 per QALY
Petrou *et al.*, 2014, Cyprus[78]	Cypriot health-care payer, 3.5%, 2012	Markov model with monthly cycles, 10 years	Patients with mRCC	Sorafenib (400 mg, twice daily) in addition to BSC compared with BSC	Medical and other pharmaceutical costs	OS and PFS (Phase III randomised clinical trial)	€102,059 per QALY

Author, year, country	Perspective, discounting and cost year	Model type, cycle length, time horizon	Patient population	Intervention/ comparator	Costs included	Outcomes (source)	Results
Mihajlovic *et al.* 2013, Serbia[79]	Serbian health-care payer, costs (3%) and effects (1.5%), 2013	Markov model with 8-week cycles, 18 cycles	Patients with mRCC whose disease had progressed on sunitinib and/or sorafenib	Everolimus second line in addition to BSC compared with BSC	Medical and other pharmaceutical costs	OS and PFS (RECORD-1 trial)	€86,978 per QALY, 95% CI €32,594–€425,258 per QALY
Casciano *et al.*, 2011, USA[80]	US health-care payer, 3%, 2010	Markov model, 8 week cycles, 6 years	Patients with RCC refractory to sunitinib treatment first line	Everolimus (10 mg, daily) compared with sorafenib (800 mg, daily)	Monitoring, blood tests, CT scans, and AE costs, GP visits, nurse, morphine and salvage therapy	Not explicitly reported but reference was made to OS (RECORD-1 trial and a Phase II single-arm sorafenib study)	US$89,160 per QALY
Hoyle *et al.*, 2010, UK[81]	UK NHS/PSS, 3.5%, 2007/8	Markov model, 6 week cycles, 10 years	Patients with advanced/metastatic RCC	Sorafenib (400 mg, twice daily) compared with BSC	PFS on sorafenib: consultant visits, blood tests and CT scans. PFS on BSC: GP visits, blood tests and CT scans. PD for both groups: GP visits, community nurse visits and pain medication	PFS and OS (Phase III randomised clinical trial)	£75,398 per QALY
Lopes *et al.*, 2012, USA[85]	US payer perspective	Excel-based cross-sectional budget impact model	Patients with advanced RCC who failed to respond or have become intolerant to sunitinib or sorafenib	Everolimus 10 mg/day	Drug costs, administration, and AE management costs	N/A	Budget impact per member per month cost was –US$0.03. Budget impact per member per year cost was –US$0.31

IFN, interferon; IL-2, interleukin 2.

Identified quality-of-life papers in renal cell carcinoma patients

Author, year, country	Population	Methods	Health states	Instrument (valuation)	Utility results
Escudier *et al.* 2009; France[107]	107 patients (54 assigned to sunitinib a.m. and 53 to sunitinib p.m.) with mRCC, aged ≥ 18 years; failure of one prior cytokine in the metastatic setting (with IFN-α monotherapy administered ≥ 4 weeks); ECOG PS of 0 or 1. QoL estimates were available for 104 patients	Patients were enrolled between July 2005 and February 2006, across 10 European participating centres. Patients completed baseline EQ-5D questionnaires. Scores were compared with the US general population normal values	mRCC with sunitinib a.m.; mRCC with sunitinib p.m. (before bed)	EQ-5D	*a.m.* EQ-5D index 0.8 (baseline); no change from baseline to cycle 29 EQ-VAS 70 (baseline); no change from baseline to cycle 29 *p.m.* EQ-5D index 0.8 (baseline); no change from baseline to cycle 29 EQ-VAS 70 (baseline); no change from baseline to cycle 29 *Between groups* No statistically significant differences between groups for either EQ-5D Index and EQ-VAS *Valuation* EQ-VAS age-matched sample (aged 55–74 years) EQ-VAS males EQ-VAS respondents with a chronic medical condition

Author, year, country	Population	Methods	Health states	Instrument (valuation)	Utility results
Motzer *et al.* 2006; USA[171]	63 patients with cytokine-refractory (IFN-α, IL-2) mRCC; ECOG PS 0 or 1. QoL estimates were available for 60 patients	Patients were enrolled between January and July 2003. Patients completed the EQ-5D questionnaire at baseline. Assessments were performed on days 1 and 28 of each cycle. Scores were compared with the US general population normal values	mRCC with sunitinib	EQ-5D	Mean EQ-VAS 77.1 (baseline); scores through 24 weeks of treatment similar to baseline Median EQ-VAS 88.0 (baseline); scores through 24 weeks of treatment similar to baseline *Valuation* Baseline EQ-VAS similar to age-matched US general population
Karakiewicz *et al.* 2016; Australia and Canada[131]	15 patients with mRCC; aged ≥ 18 years; failure of one prior systemic first-line regimen (either single agent or combination of IL-2, interferon, bevacizumab, sunitinib, pazopanib, tivozanib, temsirolimus, or everolimus); ECOG PS 0 or 1. Prior treatment regimens included sunitinib ($n = 13$), pazopanib ($n = 1$), and tivozanib ($n = 1$) therapies. QoL estimates were available for 15 patients	Patients were enrolled between March 2012 and March 2014 (5 from one centre in Canada and 10 from three centres in Australia). Patient-reported outcomes were assessed using the EQ-5D questionnaire administered on cycle 1 day 1 before dosing and before any other clinical assessments, then every 4 weeks, at end of study treatment/withdrawal, and at follow-up (28 days after last dose)	mRCC with axitinib	EQ-5D	Mean EQ-5D (baseline) 0.7947; mean change from baseline to end of treatment –0.0837 Mean EQ-VAS (baseline) 73.3; mean change from baseline to end of treatment –6.5
Cella *et al.* 2012; USA[130]	435 patients (289 assigned to pazopanib and 145 to placebo); with locally advanced and mRCC; aged ≥ 18 years; treatment naive or cytokine pretreated (78 patients on placebo and 155 on pazopanib were cytokine pretreated); ECOG PS of 0 or 1. QoL estimates were available for 398 patients	EQ-5D data were collected at baseline, weeks 6, 12, 18, 24 and 48 and period of best response (period between first date of best response and progression was considered a period of best response). The preference-based EQ-5D algorithm derived from the general population in the UK by Dolan[135] was used. Results were interpreted using previously established minimally important differences: EQ-5D utility index, minimally important difference = 0.08 and EQ-5D VAS, minimally important difference = 79	mRCC with pazopanib mRCC with placebo	EQ-5D	Placebo EQ-5D utility index: Mean ± SD at baseline: 0.73 ± 0.24 ($n = 143$) Mean ± SD at week 48: 0.80 ± 0.24 ($n = 24$) Mean change from baseline at week 48:–0.01 ± 0.20 ($n = 24$) Placebo EQ-5D VAS score: Mean ± SD at baseline: 65.9 ± 23.84 ($n = 141$) Mean ± SD at week 48: 73.1 ± 17.29 ($n = 23$) Mean change from baseline at week 48: 8.8 ± 23.96 ($n = 23$)

Author, year, country	Population	Methods	Health states	Instrument (valuation)	Utility results
					Pazopanib EQ-5D utility index: Mean ± SD at baseline: 0.72 ± 0.25 ($n = 287$) Mean ± SD at week 48: 0.79 ± 0.20 ($n = 98$) Mean change from baseline at week 48: 0.03 ± 0.20 ($n = 98$) Pazopanib EQ-5D VAS score: Mean ± SD at baseline: 64.6 ± 23.69 ($n = 283$) Mean ± SD at week 48: 72.0 ± 17.78 ($n = 95$) Mean change from baseline at week 48: 2.4 ± 24.21 ($n = 95$)
Cella *et al.* 2013; USA[67]	723 patients (361 assigned to axitinib and 362 to sorafenib) with advanced RCC after failure of one first-line systemic regimen; aged ≥ 18 years	Patient-reported outcomes were assessed using the EQ-5D and were completed at screening, after every 4 weeks of therapy, at end of study treatment, and at follow-up (28 days after end of therapy). The index scoring algorithm derived from a UK general population sample[168] was used	mRCC with axitinib mRCC with sorafenib	EQ-5D	*Axitinib* Estimated mean EQ-5D 0.71 Estimated mean EQ-5D VAS 68.11 *Sorafenib* Estimated mean EQ-5D 0.69 Estimated mean EQ-5D VAS 68.64 *Between groups* Observed EQ-5D means similar until end of treatment Estimated mean EQ-5D 0.02 (mixed-effects model, $p = 0.1903$, 95% CI –0.01 to 0.05; interaction between treatment and time, $p = 0.8048$) Estimated mean EQ-5D VAS –0.53 (mixed-effect model, $p = 0.6454$, 95% CI –2.77 to 1.72; interaction between treatment and time, $p = 0.1799$)

Author, year, country	Population	Methods	Health states	Instrument (valuation)	Utility results
Cella *et al.* 2016; USA[132]	821 patients (410 assigned to nivolumab and 411 to everolimus) with advanced RCC; aged ≥ 18 years; had received one (71% in nivolumab group; 72% in everolimus) or two (29% in nivolumab group; 28% in everolimus) anti-angiogenic therapies. QoL estimates were available for 706 patients	Patients were enrolled between Oct 2012 and March 2014 at 146 oncology centres in 24 countries in North America, Europe, Australia, South America and Asia. EQ-5D assessments were made at baseline, after randomisation but before cycle 1 of therapy, on day 1 of each cycle (starting with cycle 2), and at the first two follow-up visits. For each assessment, questionnaires were completed before physician contact, treatment dosing, or any procedures. EQ-5D assessments were also collected at follow-up visits at roughly 30 and 100 days after last dose. EQ-5D assessments were collected at each of the ten survival follow-up visits, which occurred every 3 months. EQ-5D data were analysed for all patients who underwent randomisation and had a baseline assessment and at least one post-baseline assessment. No adjustments for missing EQ-5D data were made	mRCC with nivolumab mRCC with everolimus	EQ-5D	*Nivolumab* Baseline EQ-5D utility index: 0.78 (SD 0.24), scores improved from baseline to week 104 Baseline EQ-5D VAS 73.3 (SD 18.5), scores improved from baseline to week 104 Clinically meaningful HRQoL improvement with nivolumab, 192 (53%) of 360 patients *Everolimus* Baseline EQ-5D utility index: 0.78 (SD 0.21), deterioration occurred from baseline to week 104 Baseline EQ-5D VAS 72.5 (SD 18.7), deterioration occurred from baseline to week 104 *Between groups* EQ-5D utility index, difference in mean change from baseline to end point: 0.04, 95% CI 0.02 to 0.07; $p = 0.0003$ (nivolumab compared with everolimus) EQ-5D VAS, difference in mean change from baseline to end point: 5.7, 95% CI 3.8 to 7.7; $p < 0.000$.

IFN, interferon; IL-2, interleukin 2; PS, performance status; QoL, quality of life.

Appendix 9 Quality assessment

Clinical literature

Randomised controlled trial Cochrane risk of bias

AXIS

	Risk assessment	Comments
Random sequence generation	Low risk	Randomisation lists were generated from an independent randomisation group using a permuted block design of size four (two to axitinib and two to sorafenib within each stratum)
Allocation concealment	Low risk	A web-enabled centralised registration system concealed treatment allocation before registration
Blinding (participants, personnel)	High risk	Open label, patients and investigators were not masked to study treatment, but PFS and ORR was assessed by a masked IRC
PFS	HR 0.66, 95% CI 0.552 to 0.779; $p < 0.0001$ (Motzer *et al.* 2013,[43] final data cut-off point November 2011)	
Blinding of outcome assessment	Low risk	PFS was assessed by a masked IRC
Incomplete outcome data	Low risk	PFS was based on the ITT population Number of patients lost to follow-up during the full duration of the trial was low (1 in the axitinib group, 3 in the sorafenib group) At the November 2011 data cut-off point, 67% of patients on axitinib had progressed compared with 64% of patients on sorafenib
Selective reporting	Low risk	PFS was prespecified in the methods, assessed and reported based on the ITT population
Other biases	Unclear risk	Dose increases were allowed in the axitinib arm but not in the sorafenib arm
OS	HR 0.97, 95% CI 0.800 to 1.174; $p = 0.3744$ (Motzer *et al.* 2013,[43] final data cut-off point November 2011)	
Blinding of outcome assessment	Low risk	Owing to the objective nature of the outcome the lack of blinding of outcome assessment is unlikely to bias the results
Incomplete outcome data	Low risk	OS was based on the ITT population. It was prespecified in the methods that a total of 417 events were required to detect improvement in median OS. At data cut-off point on November 2011, 425 events/deaths had occurred Number of patients lost to follow-up during the full duration of the trial was low (1 in the axitinib group, 3 in the sorafenib group)
Selective reporting	Low risk	OS was the primary outcome, it was prespecified in the methods, assessed and reported based on the ITT population

	Risk assessment	Comments
Other biases	Low risk	A majority of patients (54% on axitinib and 57% on sorafenib) received subsequent therapies. 29% of patients on axitinib and 41% of sorafenib patients received a subsequent mTOR whereas 33% of axitinib patients and 32% sorafenib patients received a subsequent TKI. The proportion of patients who received subsequent therapies were relatively balanced between groups At data cut-off point on 31 August 2010, 221/361 (61%) in the axitinib group and 256/362 (71%) in the sorafenib group had discontinued study treatment. However, the majority of patients discontinued owing to disease progression (axitinib 75%, sorafenib 70% of patients who discontinued) Dose increases were allowed in the axitinib arm, but not in the sorafenib arm
Response	ORR (complete plus partial response), stable response and progressive response (Rini 2011,[66] interim data cut-off point, August 2010)	
Blinding of outcome assessment	Low risk	Response assessment was carried out by a blinded IRC
Incomplete outcome data	Unclear risk	Response assessments were based on the ITT population. Number of patients lost to follow-up during the full duration of the trial was low (1 in the axitinib group, 3 in the sorafenib group) Response rates by an IRC took place at an earlier data cut-off point of August 2010 compared with data cut-off point for the other efficacy outcomes; PFS and OS (November 2011). Median time to response was not reported
Selective reporting	Low risk	Response assessments pre-specified in the methods (ORR, duration of response, time to response, CR, PR, SD, PD) were all assessed and reported
Other biases	N/A	
HRQoL	EQ-5D-5L and FKSI questionnaires – mean end of treatment and repeated measures mixed-effects models (Cella *et al.*,[133] data cut-off point NR)	
Blinding of outcome assessment	High risk	Self-reported FKSI-15 and FKSI-DRS and open-label study
Incomplete outcome data	Low risk	Completion rates for the FKSI-15 and FKSI-DRS questionnaires were ≥ 86% at baseline and measurements were either available or could be imputed for ≥ 90% of participants throughout the treatment and follow-up period (28 days). Data were projected using repeated measures mixed-effects and pattern-mixed models (to explore whether or not data were missing not-at-random) and showed similar results Number of patients lost to follow-up during the full duration of the trial was low (1 in the axitinib group, 3 in the sorafenib group)
Selective reporting	Low risk	The prespecified HRQoL (FKSI-15 and FKSI-DRS) was assessed and reported
Other biases	N/A	
AEs	Treatment related AEs, total and individual (Motzer *et al.* 2013,[43] final data cut-off point, November 2011)	
Blinding of outcome assessment	High risk	Safety assessment was carried out by the study investigators and the study was open label
Incomplete outcome data	Low risk	Safety assessment was based on the number of patients who received at least one dose of study medication
Selective reporting	Low risk	Prespecified safety assessments including medical history and physical examination, vital signs, laboratory assessment and grading of severity of AEs were all assessed and reported
Other biases	N/A	

CR, complete response; IRC, independent review committee; N/A, not available; NR, not reported; PR, partial response.

Checkmate 025

	Risk assessment	Comments
Random sequence generation	Low risk	Randomisation in a 1 : 1 ratio via IVRS, with a block size of 4 and stratified by region, MSKCC risk group, and the number of previous systemic therapies (one or two) for advanced RCC
Allocation concealment	Low risk	Treatment allocation was concealed using an IVRS
Blinding (participants, personnel)	High risk	Open label, patients and investigators were not masked to study treatment
PFS	HR 0.88, 95% CI 0.75 to 1.03; $p = 0.11$ (Motzer *et al.*, 2015,[54] final data cut-off point, June 2015)	
Blinding of outcome assessment	High risk	Disease progression assessed by unblinded investigators
Incomplete outcome data	Low risk	PFS was based on the ITT population The number of patients lost to follow-up was not reported At the June 2015 data cut-off point, there were 318/410 (78%) progressed events for patients on nivolumab and 322/411 (78%) events for patients on everolimus
Selective reporting	Low risk	PFS was a prespecified outcome in the methods and was assessed and reported
Other biases	Unclear risk	Dose modifications were not permitted for nivolumab but were permitted for everolimus; 102/397 everolimus patients had at least 1 dose reduction
OS	HR 0.73, 98.5% CI 0.57 to 0.93; $p = 0.002$ (Motzer *et al.*, 2015,[54] final data cut-off point, June 2015)	
Blinding of outcome assessment	Low risk	Owing to the objective nature of the outcome the lack of blinding of outcome assessment is unlikely to bias the results
Incomplete outcome data	Unclear risk	OS was based on the ITT population This result is based on a planned interim analysis, conducted after 398 (70%) of the 569 deaths required for the final analysis had occurred. 183 (45%) deaths were in the nivolumab group compared with 215 (52%) in the everolimus group. The number of deaths indicates immature data as the number of deaths has merely reached approximately 50%
Selective reporting	Low risk	OS was the primary outcome, it was prespecified in the methods, assessed and reported
Other biases	Low risk	Dose modifications were not permitted for nivolumab but were permitted for everolimus; 102/397 everolimus patients had at least 1 dose reduction A similar proportion of patients discontinued therapy in both groups; 339/410 (82.7%) patients discontinued nivolumab treatment and 369/411 (89.8%) discontinued everolimus treatment. However, the majority of patients discontinued due to disease progression (nivolumab 70%, everolimus 69% of patients who discontinued) At the data cut-off point 67/406 (17%) patients on nivolumab group and 28/397 (7%) on everolimus continued to received treatment 55% patients who received nivolumab and 63% who received everolimus had subsequent systemic therapy. The most common after nivolumab was: everolimus 26%, axitinib 24%. Most common treatments after everolimus were axitinib 36% and pazopanib 16%

	Risk assessment	Comments
Response	ORR (complete plus stable), stable, progressive (Motzer *et al.*, 2015,[54] final data cut-off point, June 2015)	
Blinding of outcome assessment	High risk	Disease response assessed by unblinded investigators
Incomplete outcome data	Low risk	Response was based on the ITT population. The number of patients lost to follow-up was not reported The median time to response was 3.5 months for patients on nivolumab and 3.7 months for patients on everolimus. The minimum follow-up was 14 months
Selective reporting	Low risk	Response assessments prespecified in the methods were all assessed and reported
Other biases	N/A	
HRQoL	FKSI median change from baseline to week 104 (Motzer *et al.*, 2015,[54] final data cut-off point, June 2015)	
Blinding of outcome assessment	High risk	Self-reported FKSI-DRS and EQ-5D questionnaires and open-label study
Incomplete outcome data	Low risk	The FKSI-DRS questionnaire completion rate was ≥ 80% for both groups throughout the first year of the study, but at times falling to 71% for nivolumab and 60% for everolimus during the remaining treatment period. Most discontinuation/missing data due to disease progression Baseline measurements were available for between 86% and 89% of the treated populations, and repeated measures mixed-effects models were used to prorate missing values for the FKSI-DRS (but not for EQ-5D). Pattern-mixed models were also used as a SA to test the missing not at random assumption
Selective reporting	Low risk	The prespecified HRQoL (FKSI-DRS) was assessed and reported as planned (dichotomised to give the number of patients with a 2-point change). Median change, completion rates and post hoc analyses (time to improvement, mixed model analyses, EQ-5D analyses, exploratory associations with survival) also reported
Other biases	N/A	
AEs	Total and treatment related AEs (Motzer *et al.*, 2015,[54] final data cut-off point, June 2015)	
Blinding of outcome assessment	High risk	Open label, investigators assessed safety based on laboratory assessments
Incomplete outcome data	Low risk	Safety assessment was based on the number of patients who received at least one dose of study medication
Selective reporting	Low risk	Prespecified safety assessments of AEs were assessed and reported
Other biases	Unclear risk	Dose modifications were not permitted for nivolumab but were permitted for everolimus; 102/397 everolimus patients had at least 1 dose reduction

IVRS, interactive voice response system; N/A, not applicable.

METEOR

	Risk assessment	Comments
Random sequence generation	Low risk	Study treatment was assigned centrally with an IVRS. Stratified permuted blocks were used as the randomisation schema. Randomisation was stratified by the number of previous VEGFR-TKI treatments (1 or ≥ 2) and MSKCC risk group (favourable, intermediate or poor) for previously treated patients
Allocation concealment	Low risk	Study treatment was assigned centrally with an IVRS. Study personnel did not have access to the master list of blocks or block sizes
Blinding (participants, personnel)	High risk	Patients and investigators were not masked to study treatment to allow appropriate management of AEs. IRC assessed PFS and Response at data cut-off point May 2015
PFS	HR 0.51, 95% CI 0.41 to 0.62; $p < 0.0001$ (Choueiri *et al.*, 2016,[57] interim data cut-off point 22 May 2015)	
Blinding of outcome assessment	Low risk	Progression was assessed by a masked centralised IRC
Incomplete outcome data	Unclear risk	PFS data were presented for the ITT population. Number of patients lost to follow-up were not reported. PFS was assessed at an earlier data cut-off point (May 2015) compared with OS (December 2015 data cut-off point) At the data cut-off point May 2015, there were 180 (55%) progressed events for patients on cabozantinib compared with 214 (65%) events for patients on everolimus. The number of progressed events indicates data are fairly immature with approximately 50% of patients reaching progression
Selective reporting	Low risk	The primary end point was PFS assessment of the first 375 randomised patients. The ITT population of 658 patients was also reported for PFS
Other biases	N/A	
OS	HR 0.66, 95% CI 0.53 to 0.83; $p = 0.00026$ (Choueiri *et al.*, 2016,[57] final data cut-off point of December 2015)	
Blinding of outcome assessment	Low risk	Owing to the objective nature of the outcome the lack of blinding of outcome assessment is unlikely to bias the results
Incomplete outcome data	Unclear risk	OS data were presented for the ITT population for an unplanned second interim analysis at the December 2015 data cut-off point. At this analysis, a total of 320 deaths occurred, 78% of the 408 deaths planned for the final OS analysis. 140 (42%) deaths were in the cabozantinib group compared with 180 (55%) in the everolimus group. The number of deaths indicates immature data as the number of events has not yet reached 50% for cabozantinib Number of patients lost to follow-up was not reported
Selective reporting	Low risk	OS was prespecified in the methods, assessed and reported

	Risk assessment	Comments
Other biases	Unclear risk	Large and uneven proportions of patients discontinued therapy in the two groups; 77.9% of patients discontinued cabozantinib treatment and 92.8% discontinued everolimus treatment. However, the majority of patients discontinued due to disease progression (cabozantinib 61.9%, everolimus 64.0% of patients who discontinued) Majority of patients received subsequent treatments. The most common for both the cabozantinib and everolimus groups was axitinib (50% and 55%, respectively). The proportion of patients who received subsequent treatments was fairly balanced; 55% of patients on cabozantinib compared with 50% of patients on everolimus Patients were allowed to continue study treatment beyond radiographic progression at the discretion of the investigator. This was for patients on either cabozantinib or everolimus
Response	ORR (complete plus partial), stable and progressive response (Choueiri *et al.*, 2016,[57] interim data cut-off point May 2015)	
Blinding of outcome assessment	Low risk	Tumour response was assessed by a masked centralised IRC
Incomplete outcome data	Unclear risk	Response data were presented for the ITT population. Response outcomes were assessed at an earlier data cut-off point May 2015 compared with OS (data cut-off point December 2015). Median time to response was not reported
Selective reporting	Low risk	Response assessments pre specified in the methods were all assessed and reported
Other biases	N/A	
HRQoL	Mean change from baseline scores EQ-5D-5L and EQ-VAS and FKSI-19 (Cabozantinib company STA submission, 2016[28])	
Blinding of outcome assessment	High risk	Self-reported FKSI-19 and EQ-5D-5L questionnaires and open-label study
Incomplete outcome data	Unclear risk	Completion rates were ≥ 75% for all measures. The number of patients lost to follow-up not reported. Analyses were conducted using repeated measures mixed-effects models to impute missing values, but there is no description of SAs to test the missing at random assumption
Selective reporting	Low risk	The HRQoL measures (FKSI-19, EQ-5D-5L and EQ-VAS) were prespecified and reported in a NICE HTA submission
Other biases	N/A	
AEs	Total AE and individual events (Choueiri *et al.*, 2016,[57] final data cut-off point December 2015)	
Blinding of outcome assessment	High risk	Blinding of outcome assessment of AEs not described
Incomplete outcome data	Low risk	Safety analyses were limited to patients who received any amount of study treatment and analysed per protocol
Selective reporting	Low risk	Comprehensive AEs data reported
Other biases	N/A	

IRC, independent review committee; IVRS, interactive voice response system; N/A, not applicable; VEGFR-TKI, vascular endothelial growth factor receptor tyrosine kinase inhibitor.

RECORD-1

	Risk assessment	Comments
Random sequence generation	Low risk	Randomisation was carried out centrally via a computer system. Patients were stratified by MSKCC status and number of prior VEGFR-TKI therapies
Allocation concealment	Low risk	Double-blind, concealment of treatment allocation was via a central IVRS
Blinding (participants, personnel)	Low risk	Double-blind study design Once patients had reached progression, assessed by investigators, the patients were offered open-label everolimus
PFS	HR 0.33, 95% CI 0.25 to 0.43; $p < 0.001$ (Motzer *et al.*, 2010,[93] final data cut-off point, February 2008)	
Blinding of outcome assessment	Low risk	Double-blind with independent review panel assessment
Incomplete outcome data	Low risk	All randomised patients were assessed for PFS using an ITT analysis Patients lost to follow-up: 4 in the everolimus group and 0 in the placebo group. At the February 2008 data cut-off point 137/277 (49%) patients on everolimus and 124/139 (89%) patients on placebo had progressed
Selective reporting	Low risk	Primary outcome, prespecified in methods, was assessed and reported
Other biases	Unclear risk	Placebo patients could crossover to receive open-label everolimus. 79 patients crossed over, of which 60 had progressed within 8 weeks of enrolment. 106/139 (76.2%) patients had crossed over to everolimus before ending the double-blind treatment. The 60 patients who progressed quickly suggest these may not be highly representative of the population
OS	HR 0.60, 95% CI 0.22 to 1.65 (Korhonen 2012,[65] crossover-adjusted analysis, data cut-off point, November 2008)	
Blinding of outcome assessment	Low risk	Double-blind design until progression when patients on placebo could switch to open-label everolimus. Hence the data include patients who switched from placebo to open-label everolimus. However, owing to the objective nature of the outcome the lack of blinding of outcome assessment is unlikely to bias the results
Incomplete outcome data	Low risk	All randomised patients were assessed for OS Patients lost to follow-up: 4 in the everolimus group and 0 in the placebo group. The number of deaths at the November 2008 data cut-off point were not reported
Selective reporting	Low risk	Prespecified OS was assessed and reported at different data cut-off points
Other biases	Unclear risk	The number of patients discontinued is not reported for the November 2008 data cut-off point. Discontinuation was reported for an earlier data cut-off point, February 2008. 202/277 (72%) patients on everolimus and 133/139 (96%) on placebo had discontinued treatment. The majority of patients discontinued as a result of disease progression (everolimus 49% and placebo 89%) Placebo patients could crossover to receive open-label everolimus. 79 patients crossed over, of whom 60 had progressed within 8 weeks of enrolment. 106/139 (76.2%) patients had crossed over to everolimus before ending the double-blind treatment

	Risk assessment	Comments
		At the end of the double-blind phase, which terminated at February 2008 data cut-off point, patients who were randomised to placebo were allowed to cross over to open-label everolimus. Six patients remained on placebo at the end of double-blind and were offered open-label everolimus Crossover-corrected results using the RPSFTM are reported to account for placebo patients who received open-label everolimus. This method of crossover correction for OS was a suitable method to remove any bias due to the crossover procedure No details about other subsequent therapies are reported
Response	ORR (complete plus partial), stable and progressive disease (Motzer *et al.*, 2008,[64] interim data cut-off point, October 2007) (Motzer *et al.*, 2010,[93] final data cut-off point, February 2008)	
Blinding of outcome assessment	Low risk	Double-blind with independent review panel assessment
Incomplete outcome data	Low risk	All randomised patients were assessed for response, using an ITT analysis. Patients lost to follow-up was low: 4 in everolimus group and 0 in placebo group. Median time to response was not reported
Selective reporting	Unclear risk	Progressive response was measured at an earlier data cut-off point (October 2007) compared with ORR and stable disease outcomes (February 2008)
Other biases	Unclear risk	Placebo patients could cross over to receive open-label everolimus. 79 patients crossed over, of whom 60 had progressed within 8 weeks of enrolment. 106/139 (76.2%) patients had crossed over to everolimus before ending the double-blind treatment
HRQoL	HRs for EORTC QLQ-C30 and FKSI-DRS (Beaumont 2011,[70] final data cut-off point, February 2008; Motzer 2008,[64] interim data cut-off point, October 2007)	
Blinding of outcome assessment	High risk	Self-reported questionnaires FKSI-DRS and EORTC QLQ-C30
Incomplete outcome data	Unclear risk	HRQoL data were collected only during the double-blind part of the study. 251/277 of patients in the everolimus group and 132/139 in the placebo group had at least one HRQoL assessment Completion rates were 87% and 92% for everolimus and placebo, respectively, at baseline. These rates declined to 60% at 3-month assessment and 32% at 6-month assessment Patients lost to follow-up: 4 in everolimus group and 0 in placebo group
Selective reporting	High risk	Data were not reported for the global quality of life subscale of EORTC QLQ-C30 but no other subscale data were reported
Other biases	N/A	
AEs	Total and individual treatment-related AEs (Motzer *et al.*, 2010,[93] final data cut-off point, February 2008)	
Blinding of outcome assessment	Low risk	Double-blind design data collected until the end of the double-blind treatment procedure
Incomplete outcome data	Low risk	Safety assessment included all patients who received at least one dose of study medication
Selective reporting	Low risk	Prespecified safety assessments and grading of severity of AEs
Other biases	N/A	

IVRS, interactive voice response system; VEGFR-TKI, vascular endothelial growth factor receptor tyrosine kinase inhibitor.

ROBINS observational studies

In the following tables, N/A refers to not applicable, N to no, NI to no information, PN to probably no, PY to probably yes and Y to yes.

Calvani *et al.*[58]

Outcome

Progression-free survival; benefit of intervention.

(OS reported but not estimable in the paper and, even so, only calculated for first + second TKI combined.)

Numerical result being assessed

Median PFS sorafenib–sunitinib 11 months, sunitinib–sorafenib 3 months, HR 0.46, 95% CI 0.16 to 0.95; $p = 0.0377$.

Responses underlined are potential markers for low risk of bias, and responses in **bold** are potential markers for a risk of bias. When questions relate only to sign posts to other questions, no formatting is used.

Signalling questions		
Bias due to confounding	**Description**	**Response options**
1.1 Is there potential for confounding of the effect of intervention in this study? If N/PN to 1.1: the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered	Retrospective cohort study. The study aim was to investigate the sequence of sorafenib and sunitinib and do so by retrospective review, not by randomised comparison	**Y**
If **Y/PY** to 1.1: determine whether there is a need to assess time-varying confounding:		
1.2. Was the analysis based on splitting participants' follow-up time according to intervention received? If N/PN, answer questions relating to baseline confounding (1.4 to 1.6) If Y/PY, go to question 1.3	Analyses were carried out retrospectively based on treatment received irrespective of follow-up time Confounding may be different between the study baseline and the start of the second treatment; however, there is no time-varying for the purposes of our review since we are only concerned with the second treatment	PN
1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome? If N/PN, answer questions relating to baseline confounding (1.4 to 1.6) If Y/PY, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8)	N/A	N/A

Signalling questions		
Bias due to confounding	**Description**	**Response options**
Questions relating to baseline confounding only		
1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains?	No description of adjustments for confounds until the discussion which states *the difference in total PFS was maintained even after other factors, including age, histology, performance status, MSKCC prognostic score, number of metastatic sites or line of treatment were taken into consideration* Not clear how this was done and only age is discussed in the results. Physician's initial choice of drug, and hence the group allocations, could have been influenced by any number of factors not captured in the baseline characteristics	**PN**
1.5. If Y/PY to 1.4: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?	Confounding domains were not controlled for	N/A
1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention?	Not reported	NI
Questions relating to baseline and time-varying confounding		
1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding?	N/A	N/A
1.8. If Y/PY to 1.7: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?	N/A	N/A
Risk-of-bias judgement	No adjustments for baseline confounding although only reported to affect age	**Serious**
Optional: What is the predicted direction of bias due to confounding?	Patients were older in the sorafenib–sunitinib group ($p = 0.0429$) and more had 2 + metastatic sites ($p = 0.08$) Discussion suggests the confounding does not bias the effect in favour of one treatment sequence over the other, although the direction of the baseline imbalances and the length of PFS on the first treatment are in favour of sunitinib–sorafenib ('the difference in total PFS was maintained even after other factors, including age, histology, performance status, MSKCC prognostic score, number of metastatic sites or line of treatment were taken into consideration'). This may mean the benefit of sorafenib–sunitinib is underestimated	Favours sorafenib–sunitinib (sunitinib given as second TKI)

Bias in selection of participants into the study		
2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention? If N/PN to 2.1: go to 2.4	Retrospective study of patients who had experienced disease progression or unacceptable toxicity after first TKI (sunitinib or sorafenib) and then switched to the reciprocal TKI	**Y**
2.2. If **Y/PY** to 2.1: Were the post-intervention variables that influenced selection likely to be associated with intervention?	Patient records were chosen based on what treatments they had already received	**YPY**
2.3 If **Y/PY** to 2.2: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome?		
2.4. Do start of follow-up and start of intervention coincide for most participants?	Start of second TKI was start of assessment for PFS	Y
2.5. If **Y/PY** to 2.2 and 2.3, or **N/PN** to 2.4: Were adjustment techniques used that are likely to correct for the presence of selection biases?	No adjustments reported	NI
Risk-of-bias judgement	Retrospective study with selection based on receiving the interventions under investigation	**Serious**
Optional: What is the predicted direction of bias due to selection of participants into the study?		

Bias in classification of interventions		
3.1 Were intervention groups clearly defined?		Y
3.2 Was the information used to define intervention groups recorded at the start of the intervention?	Not stated explicitly how interventions were recorded but this was done at the time of treatment and coded/reviewed retrospectively	Y
3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome?		PN
Risk-of-bias judgement	Interventions clearly defined and specified for each group	Low
Optional: What is the predicted direction of bias due to classification of interventions?	N/A – rated as being at low risk of bias	N/A

Bias due to deviations from intended interventions		
If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2		
4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice?	Dose modifications/discontinuations decided by investigator as would be expected in clinical practice and patients lost to follow-up censored as would be done in a good RCT	PN
4.2. If **Y/PY** to 4.1: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome?		N/A

Bias due to deviations from intended interventions		
If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6		
4.3. Were important co-interventions balanced across intervention groups?		N/A
4.4. Was the intervention implemented successfully for most participants?		N/A
4.5. Did study participants adhere to the assigned intervention regimen?		N/A
4.6. If **N/PN** to 4.3, 4.4 or 4.5: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention?		N/A
Risk-of-bias judgement		Low
Optional: What is the predicted direction of bias due to deviations from the intended interventions?	N/A – rated as being at low risk of bias	N/A

Bias due to missing data		
5.1 Were outcome data available for all, or nearly all, participants?	15 and 18 people were included in the sorafenib–sunitinib and sunitinib–sorafenib groups, respectively, of which 20% (n =3) and 17% (n =3) were lost to follow-up. It is unclear what this means in the context of a retrospective study Three additional patients had not progressed on the second drug in sorafenib–sunitinib and two patients in the sunitinib–sorafenib group, and so were censored in the PFS analysis	**N**
5.2 Were participants excluded due to missing data on intervention status?	N/A as retrospective study	PN
5.3 Were participants excluded due to missing data on other variables needed for the analysis?	None reported	PN
5.4 If **PN/N** to 5.1, or **Y/PY** to 5.2 or 5.3: Are the proportion of participants and reasons for missing data similar across interventions?	See numbers recorded for 5.1	Y
5.5 If **PN/N** to 5.1, or **Y/PY** to 5.2 or 5.3: Is there evidence that results were robust to the presence of missing data?	No evidence of SAs or imputation but the study censored those who had not progressed	PY
Risk-of-bias judgement		Moderate
Optional: What is the predicted direction of bias due to missing data?	Although there were some missing data, the number is balanced so unlikely to be in a particular direction	Unpredictable

Bias in measurement of outcomes		
6.1 Could the outcome measure have been influenced by knowledge of the intervention received?	*Dose reduction, delays or discontinuation were determined independently by each investigator as well as timing of follow-up visits and evaluation of disease response* 'Disease response' was assessed according to the Response Evaluation Criteria in Solid Tumours, which is not wholly protected from measurement bias	**PY**

Bias in measurement of outcomes		
6.2 Were outcome assessors aware of the intervention received by study participants?		**PY**
6.3 Were the methods of outcome assessment comparable across intervention groups?		PY
6.4 Were any systematic errors in measurement of the outcome related to intervention received?		NI
Risk-of-bias judgement		**Serious**
Optional: What is the predicted direction of bias due to measurement of outcomes?		Unpredictable

Bias in selection of the reported result		
Is the reported effect estimate likely to be selected, on the basis of the results, from . . .		
7.1. . . . . multiple outcome *measurements* within the outcome domain?	Unclear how many assessments for PFS were taken	NI
7.2. . . . . multiple *analyses* of the intervention–outcome relationship?	Uncertain whether or not, and how, analyses were adjusted for confounders and what the analysis presented takes into account, as this is only mentioned in the discussion	**PY**
7.3. . . . . different *subgroups*?	Full trial population used in analysis	N
Risk-of-bias judgement		Moderate
Optional: What is the predicted direction of bias due to selection of the reported result?		Unpredictable

Overall bias		
Risk-of-bias judgement		**Serious**
Optional: What is the overall predicted direction of bias for this outcome?	It was not always clear whether or not the presence of bias is likely to have been in a particular direction, except for confounding which may have favoured the sorafenib–sunitinib group (sunitinib given as second TKI). The results are in favour of sunitinib–sorafenib so the benefit may be underestimated (although the authors state adjusting for baseline differences did not change the results)	May favour sorafenib–sunitinib (but unclear in some cases)

ESPN

Outcome

Median PFS – proposed benefit of the intervention.

Numerical result being assessed

The PFS median months in second-line therapy (95% CI): sunitinib 1.8 (1.4 to 10.6); everolimus 2.8 (1.4 to not available). sunitinib vs. everolimus; $p = 0.6$. Represented on KM curve figure 2B.

Responses underlined are potential markers for low risk of bias, and responses in **bold** are potential markers for a risk of bias. When questions relate only to sign posts to other questions, no formatting is used.

Signalling questions		
Bias due to confounding	**Description**	**Response options**
1.1 Is there potential for confounding of the effect of intervention in this study? If N/PN to 1.1: the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered	Patients were initially randomised to treatments and crossed over to the second treatment at progression. Second-line post crossover of interest to this review, which was not re-randomised	**Y**
If **Y/PY** to 1.1: determine whether there is a need to assess time-varying confounding:		
1.2. Was the analysis based on splitting participants' follow-up time according to intervention received? If N/PN, answer questions relating to baseline confounding (1.4 to 1.6) If Y/PY, go to question 1.3	Analysis based on treatment received irrespective of follow-up time. While a subset of randomised patients received two treatments in a sequence, we are only interested in the second-line treatment	N
1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome? If N/PN, answer questions relating to baseline confounding (1.4 to 1.6) If Y/PY, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8)	N/A	N/A
Questions relating to baseline confounding only		
1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains?	The study does not describe any methods to control for confounding domains. Exploratory analyses were only conducted for histology, and none of the other potential confounders was adjusted for, and we do not know how characteristics were distributed at the start of the second treatment	**PN**
1.5. If Y/PY to 1.4: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		N/A
1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention?		NI
Questions relating to baseline and time-varying confounding		
1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding?	N/A	N/A
1.8. If Y/PY to 1.7: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?	N/A	N/A

Questions relating to baseline and time-varying confounding		
Risk-of-bias judgement	While baseline characteristics appear fairly well balanced at the beginning of the first treatment (which was randomised), we do not know about the distribution of confounders at the beginning of the second treatment, which included a subset of those randomised (55% of arm 1 and 67% of arm 2)	**Serious**
Optional: What is the predicted direction of bias due to confounding?	Everyone who received everolimus as second-line therapy had received sunitinib as first line which had proven more effective for PFS. Although at the point of crossover all patients had progressed or had unacceptable toxicity, more patients in the second-line everolimus group had either PR or stable disease with the prior treatment (74%) compared with second-line sunitinib (62%)	Favours everolimus (as second line)

PR, partial response.

Bias in selection of participants into the study		
2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention?	Selection into the second treatment was based on progression on the *first* treatment, not characteristics observed during the treatment of interest	PN
If N/PN to 2.1: go to 2.4		
2.2. If **Y/PY** to 2.1: Were the post-intervention variables that influenced selection likely to be associated with intervention?		N/A N/A
2.3 If **Y/PY** to 2.2: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome?		
2.4. Do start of follow-up and start of intervention coincide for most participants?	Start of intervention was start of assessment for second-line PFS	Y
2.5. If **Y/PY** to 2.2 and 2.3, or **N/PN** to 2.4: Were adjustment techniques used that are likely to correct for the presence of selection biases?	No adjustments reported. The main focus of the study was on the first-line therapy	NI
Risk-of-bias judgement	Selection based on discontinuations from initial randomised treatment	Moderate
Optional: What is the predicted direction of bias due to selection of participants into the study?		Unpredictable

Bias in classification of interventions		
3.1 Were intervention groups clearly defined?		Y
3.2 Was the information used to define intervention groups recorded at the start of the intervention?		Y
3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome?	Interventions clearly defined and specified for each group	PN
Risk-of-bias judgement		Low
Optional: What is the predicted direction of bias due to classification of interventions?	N/A – rated as being at low risk of bias	N/A

Bias due to deviations from intended interventions		
If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2		
4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice?	No deviations reported and not anticipated to be different from usual practice based on study methods	PN
4.2. If **Y/PY** to 4.1: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome?		N/A
If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6		
4.3. Were important co-interventions balanced across intervention groups?		N/A
4.4. Was the intervention implemented successfully for most participants?		N/A
4.5. Did study participants adhere to the assigned intervention regimen?		N/A
4.6. If **N/PN** to 4.3, 4.4 or 4.5: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention?		N/A
Risk-of-bias judgement		Low
Optional: What is the predicted direction of bias due to deviations from the intended interventions?	N/A – rated as being at low risk of bias	N/A

Bias due to missing data		
5.1 Were outcome data available for all, or nearly all, participants?	Full consort diagram provided in the main publication shows patient flow and analysis populations for first-line and second-line therapies Figure states that all patients who received sunitinib as second line ($n = 21$) and all who received everolimus as second line ($n = 23$) were available for the crossover end point. Patients all included in an ITT analysis	Y
5.2 Were participants excluded due to missing data on intervention status?	Patients censored at date of last follow-up if missing data	N
5.3 Were participants excluded due to missing data on other variables needed for the analysis?		NI
5.4 If **PN/N** to 5.1, or **Y/PY** to 5.2 or 5.3: Are the proportion of participants and reasons for missing data similar across interventions?		N/A
5.5 If **PN/N** to 5.1, or **Y/PY** to 5.2 or 5.3: Is there evidence that results were robust to the presence of missing data?		N/A
Risk-of-bias judgement		Low
Optional: What is the predicted direction of bias due to missing data?	N/A – rated as being at low risk of bias	N/A

Bias in measurement of outcomes		
6.1 Could the outcome measure have been influenced by knowledge of the intervention received?	Independent review panel assessment for PFS	N
6.2 Were outcome assessors aware of the intervention received by study participants?	*An independent radiology panel assessed tumour response using Response Evaluation Criteria in Solid Tumours v.1.0, at 6 wk, at 12 wk, and every 12 wk thereafter* Unclear if they were aware of study drug assignment	NI
6.3 Were the methods of outcome assessment comparable across intervention groups?	RECIST criteria used for PFS assessment and the same review panel	Y
6.4 Were any systematic errors in measurement of the outcome related to intervention received?		NI
Risk-of-bias judgement	While PFS may be subject to bias, an independent radiology panel was used to prevent this	Low
Optional: What is the predicted direction of bias due to measurement of outcomes?	N/A – rated as being at low risk of bias	N/A

Bias in selection of the reported result		
Is the reported effect estimate likely to be selected, on the basis of the results, from . . .		
7.1. . . . . multiple outcome *measurements* within the outcome domain?	PFS clearly defined and reported for first- and second-line therapy. Measured by independent radiology panel	PN
7.2. . . . . multiple *analyses* of the intervention–outcome relationship?	Not apparent	PN
7.3. . . . . different *subgroups*?		N
Risk-of-bias judgement		Low
Optional: What is the predicted direction of bias due to selection of the reported result?	N/A – rated as being at low risk of bias	N/A

Overall bias		
Risk-of-bias judgement	The study is considered low risk for all domains except confounding, which is a serious risk of bias. While the study has generally controlled well for biases, our inclusion of only the second-line therapy means there is some risk of confounding from the first treatment allocation and other confounders not controlled for in the analysis	**Serious**
Optional: What is the overall predicted direction of bias for this outcome?		Unpredictable

Iacovelli *et al.*[59]

Outcome

Overall survival – proposed benefit of the intervention.

Numerical result being assessed

Overall survival multivariate cox regression adjusted for Heng prognostic criteria (univariate also available): HR 2.21, 95% CI 1.47 to 3.31; $p < 0.001$.

Responses underlined are potential markers for low risk of bias, and responses in **bold** are potential markers for a risk of bias. When questions relate only to sign posts to other questions, no formatting is used.

Signalling questions		
Bias due to confounding	**Description**	**Response options**
1.1 Is there potential for confounding of the effect of intervention in this study? If N/PN to 1.1: the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered	Retrospective cohort study to investigate the effect of bone metastases on prognosis and effectiveness of everolimus and sorafenib	**Y**
If **Y/PY** to 1.1: determine whether there is a need to assess time-varying confounding:		
1.2. Was the analysis based on splitting participants' follow-up time according to intervention received? If N/PN, answer questions relating to baseline confounding (1.4 to 1.6) If Y/PY, go to question 1.3	*Treatments were administered until disease progression or until the patient developed unacceptable levels of toxicity* Patients were reviewed retrospectively by presence of bone metastases and treatment received	N
1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome? If N/PN, answer questions relating to baseline confounding (1.4 to 1.6) If Y/PY, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8)	N/A	N/A
Questions relating to baseline confounding only		
1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains?	Only Heng score was included in the multivariate analysis. No other confounders were controlled for	**PN**
1.5. If Y/PY to 1.4: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?	Yes for prognosis using the Heng score	N/A
1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention?		NI

Signalling questions		
Bias due to confounding	**Description**	**Response options**
Questions relating to baseline and time-varying confounding		
1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding?		N/A
1.8. If Y/PY to 1.7: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		N/A
Risk-of-bias judgement	There were several possible confounders for the therapy comparison we are interested in that were not controlled for in the study's design	**Serious**
Optional: What is the predicted direction of bias due to confounding?	Unknown. Patient characteristics are presented for all patients and broken down by presence of bone metastases. The paper does present the presence of bone metastases by treatment, which was very similar for patients who received everolimus and those who had sorafenib (33%). This does not indicate bias in a particular direction	Unpredictable

Bias in selection of participants into the study		
2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention?	Inclusion criteria are well defined but patients were reviewed retrospectively so had already received the treatments of interest to this review when they were selected	**Y**
If N/PN to 2.1: go to 2.4	'We retrospectively reviewed consecutive patients with clear-cell mRCC treated with three lines of targeted therapies at 23 centres in Italy'	
2.2. If **Y/PY** to 2.1: Were the post-intervention variables that influenced selection likely to be associated with intervention?	Inclusion criteria and baseline characteristics could have influenced clinician's decisions about which therapy was given	**PY**
2.3 If **Y/PY** to 2.2: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome?		**PY**
2.4. Do start of follow-up and start of intervention coincide for most participants?	Start of intervention was start of assessment for OS	Y
2.5. If **Y/PY** to 2.2 and 2.3, or **N/PN** to 2.4: Were adjustment techniques used that are likely to correct for the presence of selection biases?	No adjustments reported	NI
Risk-of-bias judgement	Retrospective study with selection based on receiving the interventions under investigation	**Serious**
Optional: What is the predicted direction of bias due to selection of participants into the study?		Unpredictable

Bias in classification of interventions		
3.1 Were intervention groups clearly defined?	*All patients received standard dose EV or SO after two previous lines of targeted therapies; treatments were administered until disease progression or until the patient developed unacceptable levels of toxicity*	Y
3.2 Was the information used to define intervention groups recorded at the start of the intervention?	Not stated explicitly how interventions were recorded but this would have been done at the time of treatment and reviewed retrospectively	Y
3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome?		PN
Risk-of-bias judgement	Interventions were clearly defined for each therapy group, although this was not the main focus of the study (presence of bone metastases)	Low
Optional: What is the predicted direction of bias due to classification of interventions?	N/A – rated as being at low risk of bias	N/A

Bias due to deviations from intended interventions		
If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2		
4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice?	Participants were not allocated, but chosen on the basis of what they had already received in usual practice. Discontinuations decided by investigator as would be expected in clinical practice and patients lost to follow-up censored as would be done in a good RCT	PN
4.2. If **Y/PY** to 4.1: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome?		N/A
If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6		
4.3. Were important co-interventions balanced across intervention groups?		N/A
4.4. Was the intervention implemented successfully for most participants?		N/A
4.5. Did study participants adhere to the assigned intervention regimen?		N/A
4.6. If **N/PN** to 4.3, 4.4 or 4.5: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention?		N/A
Risk-of-bias judgement		Low
Optional: What is the predicted direction of bias due to deviations from the intended interventions?	N/A – rated as being at low risk of bias	N/A

Bias due to missing data		
5.1 Were outcome data available for all, or nearly all, participants?	*A total of 281 mRCC patients treated with three lines of targeted therapies were screened. Of these, 233 patients received EV or SO as third-line [i.e. met the inclusion criteria] and included in the final analysis* Paper does not mention any patients being excluded or lost to follow-up	NI
5.2 Were participants excluded due to missing data on intervention status?	Retrospective study	PN
5.3 Were participants excluded due to missing data on other variables needed for the analysis?	None reported. States 233 met the criteria and were included in the analyses	NI
5.4 If **PN/N** to 5.1, or **Y/PY** to 5.2 or 5.3: Are the proportion of participants and reasons for missing data similar across interventions?		N/A
5.5 If **PN/N** to 5.1, or **Y/PY** to 5.2 or 5.3: Is there evidence that results were robust to the presence of missing data?		N/A
Risk-of-bias judgement		Low
Optional: What is the predicted direction of bias due to missing data?	N/A – rated as being at low risk of bias	N/A

Bias in measurement of outcomes		
6.1 Could the outcome measure have been influenced by knowledge of the intervention received?	Survival free from assessor bias *OS was defined as the time from start of third-line treatment to death or censored at last contact*	PN
6.2 Were outcome assessors aware of the intervention received by study participants?	Probably yes, although paper does not go into detail of how data were extracted and coded	**PY**
6.3 Were the methods of outcome assessment comparable across intervention groups?		PY
6.4 Were any systematic errors in measurement of the outcome related to intervention received?		NI
Risk-of-bias judgement	Some aspects are not well described but OS not likely to be biased by measurement	Low
Optional: What is the predicted direction of bias due to measurement of outcomes?	N/A – rated as being at low risk of bias	N/A

Bias in selection of the reported result		
Is the reported effect estimate likely to be selected, on the basis of the results, from . . .		
7.1. . . . . multiple outcome *measurements* within the outcome domain?	OS clearly defined and well reported. No time point or type of measurement issues	N
7.2. . . . . multiple *analyses* of the intervention–outcome relationship?	Several analyses were undertaken and all appear to be reported in the paper	PN
7.3. . . . . different *subgroups*?	Bone metastases, sites of metastases and prognostic subgroups are reported in the paper as planned	PN
Risk-of-bias judgement		Low
Optional: What is the predicted direction of bias due to selection of the reported result?	N/A – rated as being at low risk of bias	N/A

Overall bias		
Risk-of-bias judgement	While the study is considered to be at low risk for most domains, there are serious risks of confounding and selection biases that have not been controlled for. This is mainly because of the retrospective design of the study and its primary focus on the presence of metastases	**Serious**
Optional: What is the overall predicted direction of bias for this outcome?		Unpredictable

EV, everolimus; SO, sorafenib.

Paglino *et al.*[60]

Outcome

Progression-free survival – proposed benefit of the intervention.

Numerical result being assessed

Progression-free survival median months in second-line therapy (95% CI): sunitinib 3.90 (3.00 to 13.42); sorafenib 9.12 (3.50 to 20.03). sunitinib vs. sorafenib; $p = 0.2379$ (see KM, figure 3).

Responses <u>underlined</u> are potential markers for low risk of bias, and responses in **bold** are potential markers for a risk of bias. Where questions relate only to sign posts to other questions, no formatting is used.

Signalling questions		
Bias due to confounding	**Description**	**Response**
1.1 Is there potential for confounding of the effect of intervention in this study? If <u>N/PN</u> to 1.1: the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered	Retrospective cohort review of sorafenib-mTORi-sunitinib compared with sunitinib-mTORi-sorafenib sequencing with only the final TKI treatment of interest to this review	**Y**
If **Y/PY** to 1.1: determine whether there is a need to assess time-varying confounding:		
1.2. Was the analysis based on splitting participants' follow-up time according to intervention received? If N/PN, answer questions relating to baseline confounding (1.4 to 1.6) If Y/PY, go to question 1.3		N
1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome? If N/PN, answer questions relating to baseline confounding (1.4 to 1.6) If Y/PY, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8)	N/A	N/A

Signalling questions		
Bias due to confounding	**Description**	**Response**
Questions relating to baseline confounding only		
1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains?	*Cox's regression models were used to analyse associations between PFS and baseline characteristics and treatment groups* Rather than controlling for confounds, this was done with univariate analysis to explore the influence of each baseline characteristic. Not all confounds we highlighted are listed	**PN**
1.5. If Y/PY to 1.4: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		N/A
1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention?		NI
Questions relating to baseline and time-varying confounding		
1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding?	N/A	N/A
1.8. If Y/PY to 1.7: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?	N/A	N/A
Risk-of-bias judgement	Although explored with univariate analyses, the presence of multiple possible confounders and baseline imbalances that were not controlled for are likely to introduce a critical risk of bias	**Critical**
Optional: What is the predicted direction of bias due to confounding?	Those who received sunitinib as third line all had clear-cell histology and a smaller proportion had Fuhrman's grade 3 or 4. ECOG fairly balanced. However, more in that group also had liver metastases so not all differences favour this group	Unpredictable

Bias in selection of participants into the study		
2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention?	Yes, selection occurred after patients had received all treatments	**Y**
If N/PN to 2.1: go to 2.4		
2.2. If **Y/PY** to 2.1: Were the post-intervention variables that influenced selection likely to be associated with intervention?	Patient records were chosen based on what treatments they had already received	**Y**
2.3 If **Y/PY** to 2.2: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome?	Progression or AEs on the first therapy might be related to the likelihood of progressing on the second/third therapy	**PY**
2.4. Do start of follow-up and start of intervention coincide for most participants?		Y
2.5. If **Y/PY** to 2.2 and 2.3, or **N/PN** to 2.4: Were adjustment techniques used that are likely to correct for the presence of selection biases?	No adjustments reported. The main focus of the study was on first/second/third treatment sequence as a whole, not on the third line that we are focusing on	NI

Bias in selection of participants into the study		
Risk-of-bias judgement	The retrospective design does not control for selection biases	**Serious**
Optional: What is the predicted direction of bias due to selection of participants into the study?	Unknown	Unpredictable

Bias in classification of interventions		
3.1 Were intervention groups clearly defined?	Yes, drugs and schedules defined clearly	Y
3.2 Was the information used to define intervention groups recorded at the start of the intervention?		Y
3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome?	Not likely as done at the time treatment was given and reviewed retrospectively	PN
Risk-of-bias judgement		Low
Optional: What is the predicted direction of bias due to classification of interventions?	N/A – rated as being at low risk of bias	N/A

Bias due to deviations from intended interventions		
If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2		
4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice?	Probably not – patients included based on what treatments they had already received *Patients who were treated with any other agent during the gap between the three drugs treatments were excluded from this analysis*	PN
4.2. If **Y/PY** to 4.1: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome?		N/A
If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6		
4.3. Were important co-interventions balanced across intervention groups?		N/A
4.4. Was the intervention implemented successfully for most participants?		N/A
4.5. Did study participants adhere to the assigned intervention regimen?		N/A
4.6. If **N/PN** to 4.3, 4.4 or 4.5: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention?		N/A
Risk-of-bias judgement		Low
Optional: What is the predicted direction of bias due to deviations from the intended interventions?	N/A – rated as being at low risk of bias	N/A

Bias due to missing data		
5.1 Were outcome data available for all, or nearly all, participants?	No description of exclusions or missing data	NI
5.2 Were participants excluded due to missing data on intervention status?		NI
5.3 Were participants excluded due to missing data on other variables needed for the analysis?		NI
5.4 If **PN/N** to 5.1, or **Y/PY** to 5.2 or 5.3: Are the proportion of participants and reasons for missing data similar across interventions?		N/A
5.5 If **PN/N** to 5.1, or **Y/PY** to 5.2 or 5.3: Is there evidence that results were robust to the presence of missing data?		N/A
Risk-of-bias judgement		No information
Optional: What is the predicted direction of bias due to missing data?		N/A

Bias in measurement of outcomes		
6.1 Could the outcome measure have been influenced by knowledge of the intervention received?	Progression assessment can depend on the criteria used which may be applied subjectively if assessors had known treatments received	**PY**
6.2 Were outcome assessors aware of the intervention received by study participants?	Does not say they were independent *The status of disease progression during the three treatment periods was determined by radiological assessment using the Response Evaluation Criteria In Solid Tumours (RECIST) approximately every 12 weeks*	**PY**
6.3 Were the methods of outcome assessment comparable across intervention groups?		PY
6.4 Were any systematic errors in measurement of the outcome related to intervention received?		NI
Risk-of-bias judgement	PFS may be subject to bias in assignment and there is no description of blinded assessors	**Serious**
Optional: What is the predicted direction of bias due to measurement of outcomes?		Unpredictable

Bias in selection of the reported result		
Is the reported effect estimate likely to be selected, on the basis of the results, from . . .		
7.1. . . . multiple outcome *measurements* within the outcome domain?	PFS clearly defined and reported for first-, second- and third-line therapy, and overall	NI
7.2. . . . multiple *analyses* of the intervention–outcome relationship?	Various univariate regression analyses reported in table II	**PY**
7.3. . . . different *subgroups*?		PN
Risk-of-bias judgement		Moderate
Optional: What is the predicted direction of bias due to selection of the reported result?		Unpredictable

Overall bias	
Risk-of-bias judgement	**Serious**
Optional: What is the overall predicted direction of bias for this outcome?	Unpredictable

Porta *et al.*[61]

Outcome

Progression-free survival – proposed benefit of the intervention.

Numerical result being assessed

Progression-free survival: HR 0.535 (95% CI 0.387 to 0.740); $p = 0.0002$ (Porta 2010[153] KM data figure 1c, Porta 2011[61]). Median PFS: 7.89 months (95% CI 0.8 to 26.9) with sunitinib and 4.24 months (95% CI 0.1 to 34.7) with sorafenib.

Responses underlined are potential markers for low risk of bias, and responses in **bold** are potential markers for a risk of bias. Where questions relate only to sign posts to other questions, no formatting is used.

Signalling questions		
Bias due to confounding	**Description**	**Response**
1.1 Is there potential for confounding of the effect of intervention in this study? If N/PN to 1.1: the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered	Retrospective cohort study of sunitinib-sorafenib and sorafenib-sunitinib sequencing with only second TKI of interest to this review	**Y**
If **Y/PY** to 1.1: determine whether there is a need to assess time-varying confounding:		
1.2. Was the analysis based on splitting participants' follow-up time according to intervention received? If N/PN, answer questions relating to baseline confounding (1.4 to 1.6) If Y/PY, go to question 1.3	Analyses were carried out retrospectively based on whether patients received sunitinib then sorafenib, or sorafenib then sunitinib. All the time patients spent on each drug was taken into consideration. Initiation of the second TKI occurred when a patient had progressed on the first treatment	N
1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome? If N/PN, answer questions relating to baseline confounding (1.4 to 1.6) If Y/PY, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8)		N/A
Questions relating to baseline confounding only		
1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains?	The study tested the influence of several confounders using multivariate analyses but second TKI result not adjusted. There was an important imbalance in MSKCC and treatments between study centres	**N**

Signalling questions		
Bias due to confounding	**Description**	**Response**
1.5. If Y/PY to 1.4: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		N/A
1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention?		NI
Questions relating to baseline and time-varying confounding		
1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding?		N/A
1.8. If Y/PY to 1.7: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		N/A
Risk-of-bias judgement	The imbalance in MSKCC was not controlled for in the result we need for the review	**Critical**
Optional: What is the predicted direction of bias due to confounding?		Favours sorafenib-sunitinib (i.e. second-line sunitinib group)

Bias in selection of participants into the study		
2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention?	Eligible patients were identified from records and analysed retrospectively based on their use of both medicines	**Y**
If N/PN to 2.1: go to 2.4		
2.2. If **Y/PY** to 2.1: Were the post-intervention variables that influenced selection likely to be associated with intervention?	At the time the therapy choice was made, patient characteristics affecting choice of therapy may be related to the patient's prognosis	**PY**
2.3 If **Y/PY** to 2.2: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome?		**PY**
2.4. Do start of follow-up and start of intervention coincide for most participants?		Y
2.5. If **Y/PY** to 2.2 and 2.3, or **N/PN** to 2.4: Were adjustment techniques used that are likely to correct for the presence of selection biases?		NI
Risk-of-bias judgement	Between centres, the selection of participants was very varied in terms of which TKI was given first. This may represent a preferred sequence in given centres, or selection bias in the choice of centres or patients	**Serious**
Optional: What is the predicted direction of bias due to selection of participants into the study?		Unpredictable

Bias in classification of interventions		
3.1 Were intervention groups clearly defined?	Study report states clearly the intervention type, dose and sequence for each group	Y
3.2 Was the information used to define intervention groups recorded at the start of the intervention?	This was done at the time of treatment and reviewed retrospectively	Y
3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome?	Owing to the retrospective nature of the study, it is possible that study authors were aware of the health status of participants which could have affected classification or selection into the study	PN
Risk-of-bias judgement		Low
Optional: What is the predicted direction of bias due to classification of interventions?	N/A – rated as being at low risk of bias	N/A

Bias due to deviations from intended interventions		
If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2		
4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice?	The retrospective and real-life nature of this study means participants were not allocated, but chosen on the basis of what they had already received in usual practice *Patients who were treated with any other agent during the treatment gap between sorafenib and sunitinib therapy were excluded from the present study*	PN
4.2. If **Y/PY** to 4.1: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome?		N/A
If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6		
4.3. Were important co-interventions balanced across intervention groups?		N/A
4.4. Was the intervention implemented successfully for most participants?		N/A
4.5. Did study participants adhere to the assigned intervention regimen?		N/A
4.6. If **N/PN** to 4.3, 4.4 or 4.5: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention?		N/A
Risk-of-bias judgement		Low
Optional: What is the predicted direction of bias due to deviations from the intended interventions?	N/A – rated as being at low risk of bias	N/A

Bias due to missing data		
5.1 Were outcome data available for all, or nearly all, participants?	No mention of missing data	NI
5.2 Were participants excluded due to missing data on intervention status?		NI
5.3 Were participants excluded due to missing data on other variables needed for the analysis?	*Patients who remained on the second TKI without disease progression at the end of the study period were censored from the analysis*	NI
5.4 If **PN/N** to 5.1, or **Y/PY** to 5.2 or 5.3: Are the proportion of participants and reasons for missing data similar across interventions?		NI
5.5 If **PN/N** to 5.1, or **Y/PY** to 5.2 or 5.3: Is there evidence that results were robust to the presence of missing data?	There is no evidence of sensitivity or exploratory analyses being carried out	NI
Risk-of-bias judgement		No information
Optional: What is the predicted direction of bias due to missing data?		N/A

Bias in measurement of outcomes		
6.1 Could the outcome measure have been influenced by knowledge of the intervention received?	Study states: *The status of disease progression during the first and second TKI was determined by radiological assessment (Response Evaluation Criteria In Solid Tumours [RECIST]) approximately every 12 weeks* Measurement can be interpreted subjectively	**PY**
6.2 Were outcome assessors aware of the intervention received by study participants?	Does not say this was done by blinded assessors	**PY**
6.3 Were the methods of outcome assessment comparable across intervention groups?		NI
6.4 Were any systematic errors in measurement of the outcome related to intervention received?		NI
Risk-of-bias judgement		**Serious**
Optional: What is the predicted direction of bias due to measurement of outcomes?		Unpredictable

Bias in selection of the reported result		
Is the reported effect estimate likely to be selected, on the basis of the results, from . . .		
7.1. . . . multiple outcome *measurements* within the outcome domain?	PFS judged with RECIST	NI
7.2. . . . multiple *analyses* of the intervention–outcome relationship?	RECIST used which includes various subjective judgements. PFS during first, second and both treatments together are all reported. Study report includes information about multivariate analysis methods and results	**PY**
7.3. . . . different *subgroups*?		PN
Risk-of-bias judgement		Moderate
Optional: What is the predicted direction of bias due to selection of the reported result?		Unpredictable

Overall bias		
Risk-of-bias judgement	The critical risk of bias in this study relates to confounding as a result of a baseline imbalance in MSKCC criteria which was not controlled for in the result of interest	**Critical**
Optional: What is the overall predicted direction of bias for this outcome?	Direction of bias is mostly difficult to predict, but where there was information (such as for confounding), the bias is likely to have favoured the sorafenib–sunitinib group (second-line sunitinib)	Mostly unpredictable but may favour second-line sunitinib

SWITCH

Outcome

Progression-free survival – proposed benefit of the intervention.

Numerical result being assessed

Progression-free survival: HR 0.55 (95% CI 0.41 to 0.74); $p < 0.001$ (Eichelberg *et al.*,[56] figure 3B, p. 842).
Median PFS: sunitinib 5.4 months (95% CI 3.0 to 5.5), sorafenib 2.8 months (95% CI 2.7 to 2.9).

Responses <u>underlined</u> are potential markers for low risk of bias, and responses in **bold** are potential markers for a risk of bias. Where questions relate only to sign posts to other questions, no formatting is used.

Signalling questions		
Bias due to confounding	**Description**	**Response**
1.1 Is there potential for confounding of the effect of intervention in this study? If <u>N/PN</u> to 1.1: the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered	The study was a RCT design but only a subset of the full population initiated the second treatment. Only the final TKI treatment of interest to this review so considered observational	**Y**
If **Y/PY** to 1.1: determine whether there is a need to assess time-varying confounding:		
1.2. Was the analysis based on splitting participants' follow-up time according to intervention received? If N/PN, answer questions relating to baseline confounding (1.4 to 1.6) If Y/PY, go to question 1.3		PN
1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome? If N/PN, answer questions relating to baseline confounding (1.4 to 1.6) If Y/PY, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8)	Intervention switches were planned at the point of progression/toxicity	N/A

Signalling questions		
Bias due to confounding	**Description**	**Response**
Questions relating to baseline confounding only		
1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains?	The study was designed as a RCT but the second phase is subject to possible confounding because not all randomised patients crossed over. Baseline characteristics are available for the subset who initiated the second treatment, with no obvious imbalances, but this is not for the point they crossed over to the second treatment	**PN**
1.5. If Y/PY to 1.4: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		N/A
1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention?		NI
Questions relating to baseline and time-varying confounding		
1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding?		N/A
1.8. If Y/PY to 1.7: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?		N/A
Risk-of-bias judgement	Although most known confounding variables were balanced at the start of the RCT for the subset starting second treatment, we do not know how they compared at crossover. Additionally, patients in each group had just received the opposite treatment	**Serious**
Optional: What is the predicted direction of bias due to confounding?		Unpredictable

Bias in selection of participants into the study		
2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention?	Selection into study was not, but selection into the second treatment was based on progression on the first treatment (i.e. not the intervention of interest)	PN
If N/PN to 2.1: go to 2.4		
2.2. If **Y/PY** to 2.1: Were the post-intervention variables that influenced selection likely to be associated with intervention?	The different safety profiles of sorafenib and sunitinib may have contributed to differences in first-line therapy discontinuation, and thus affected first-line PFS and total PFS. In sequential studies, the decision to end first-line treatment can potentially be influenced by investigator knowledge that a second-line treatment is readily available. In SWITCH, however, confirmed radiologic progression was required to proceed to second-line treatment	N/A
2.3 If **Y/PY** to 2.2: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome?		N/A
2.4. Do start of follow-up and start of intervention coincide for most participants?		Y
2.5. If **Y/PY** to 2.2 and 2.3, or **N/PN** to 2.4: Were adjustment techniques used that are likely to correct for the presence of selection biases?		N/A

Bias in selection of participants into the study		
Risk-of-bias judgement	See above quote regarding initiation of the second treatment, which constitutes selection for our purposes. The use of radiological confirmation does not necessarily remove this risk as criteria can be applied subjectively	Moderate
Optional: What is the predicted direction of bias due to selection of participants into the study?	A higher proportion of patients who received first-line sunitinib were not treated with second-line sorafenib (44%) than vice versa (32%). In particular, more in the former group were not treated because of AEs. Unclear if this favours one over the other	Unpredictable

Bias in classification of interventions		
3.1 Were intervention groups clearly defined?	Study report states clearly the intervention type, dose and sequence for each group	Y
3.2 Was the information used to define intervention groups recorded at the start of the intervention?		Y
3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome?		PN
Risk-of-bias judgement		Low
Optional: What is the predicted direction of bias due to classification of interventions?	N/A – rated as being at low risk of bias	N/A

Bias due to deviations from intended interventions		
If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2		
4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice?	Consort diagram shows a small percentage in each group did not receive the intended first-line therapy – unlikely to be significantly more than expected in usual practice. Balanced (3% first-line sorafenib; 4% first-line sunitinib) *In particular, patients were censored if they received unauthorised cancer treatment (without progressive disease or other status counted as an event), which included patients who received off-protocol second-line therapy instead of per-protocol second-line therapy* *Eichelberg* et al.[56]	PN
4.2. If **Y/PY** to 4.1: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome?	Unlikely to have affected the outcome as data censored	N
If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6		
4.3. Were important co-interventions balanced across intervention groups?		N/A
4.4. Was the intervention implemented successfully for most participants?		N/A
4.5. Did study participants adhere to the assigned intervention regimen?		N/A
4.6. If **N/PN** to 4.3, 4.4 or 4.5: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention?		N/A

Bias due to deviations from intended interventions		
Risk-of-bias judgement		Low
Optional: What is the predicted direction of bias due to deviations from the intended interventions?	N/A – rated as being at low risk of bias	N/A

Bias due to missing data		
5.1 Were outcome data available for all, or nearly all, participants?	Of the patients who stopped first-line treatment, a large proportion of each group did not initiate second-line therapy (32% first-line sorafenib; 44% sunitinib – for various reasons). ITT analysis used for those who did initiate second-therapy and less than 10% lost to follow-up	PY
5.2 Were participants excluded due to missing data on intervention status?	*For time-to-event analysis, missing values were censored*	N
5.3 Were participants excluded due to missing data on other variables needed for the analysis?	ITT analysis used and all patients accounted for in the analysis	N
5.4 If **PN/N** to 5.1, or **Y/PY** to 5.2 or 5.3: Are the proportion of participants and reasons for missing data similar across interventions?		N/A
5.5 If **PN/N** to 5.1, or **Y/PY** to 5.2 or 5.3: Is there evidence that results were robust to the presence of missing data?		N/A
Risk-of-bias judgement	While there are likely to be some missing data, ITT population was used and the paper is clear about which data were censored. Reasons for discontinuations are defined	Low
Optional: What is the predicted direction of bias due to missing data?	N/A – rated as being at low risk of bias	N/A

Bias in measurement of outcomes		
6.1 Could the outcome measure have been influenced by knowledge of the intervention received?	*The study was open-label rather than double-blind, introducing a potential for investigator bias; however, the protocol mandated that confirmed radiologic progression was required to stop treatment on the grounds of disease progression, which reduced this potential for bias* However, PFS measurement can be interpreted subjectively	**PY**
6.2 Were outcome assessors aware of the intervention received by study participants?	Open-label study	**Y**
6.3 Were the methods of outcome assessment comparable across intervention groups?		PY
6.4 Were any systematic errors in measurement of the outcome related to intervention received?		NI
Risk-of-bias judgement		**Serious**
Optional: What is the predicted direction of bias due to measurement of outcomes?		Unpredictable

Bias in selection of the reported result		
Is the reported effect estimate likely to be selected, on the basis of the results, from . . .		
7.1. . . . multiple outcome *measurements* within the outcome domain?	PFS judged with RECIST which includes several factors	**PY**
7.2. . . . multiple *analyses* of the intervention–outcome relationship?	Cox PH model used. PFS during first, second and both treatments together are all reported, and stratified. Definition of analysis matches clinicaltrials.gov	PN
7.3. . . . different *subgroups*?		N
Risk-of-bias judgement		Moderate
Optional: What is the predicted direction of bias due to selection of the reported result?		Unpredictable

Overall bias		
Risk-of-bias judgement	Moderate risk relating to possible confounding, even though the first phase was randomised, and relating to measurement of PFS. Serious risks of selection bias into the second-line treatment phase	**Serious**
Optional: What is the overall predicted direction of bias for this outcome?		Unpredictable

Vogelzang *et al.*[62]

Outcome

Benefits: PFS and OS.

Numerical result being assessed

Multivariable-adjusted analysis: OS (HR 1.16, 95% CI 0.74 to 1.82) and PFS (HR 1.16, 95% CI 0.85 to 1.59), table 2, p. 744.

Responses underlined are potential markers for low risk of bias, and responses in **bold** are potential markers for a risk of bias. Where questions relate only to sign posts to other questions, no formatting is used.

Signalling questions		
Bias due to confounding	**Description**	**Response**
1.1 Is there potential for confounding of the effect of intervention in this study? If N/PN to 1.1: the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered	Retrospective chart review – not a randomised study. Patient records selected by physicians for inclusion so high risk or confounding at baseline	**Y**

Signalling questions		
Bias due to confounding	**Description**	**Response**
If **Y/PY** to 1.1: determine whether there is a need to assess time-varying confounding:		
1.2. Was the analysis based on splitting participants' follow-up time according to intervention received? If N/PN, answer questions relating to baseline confounding (1.4 to 1.6) If Y/PY, go to question 1.3	No limitation on follow-up; follow-up was decided based on individual patient and physician	N
1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome? If N/PN, answer questions relating to baseline confounding (1.4 to 1.6) If Y/PY, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8)	N/A	N/A
Questions relating to baseline confounding only		
1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains?	*Multivariable Cox proportional hazards models adjusted for age, sex, metastasized RCC at initial diagnosis, prior nephrectomy, type and duration of first targeted therapy, clinical benefit while on first targeted therapy (physician assessed yes/no), occurrence of progression while on first targeted therapy (physician assessed yes/no), duration of mRCC at second targeted therapy initiation, sites of metastases, clear-cell RCC histology, comorbid hypercholesterolemia, ECOG performance status, and years of practice of the treating physician*	PY
1.5. If Y/PY to 1.4: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?	Variables were baseline characteristics	PY
1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention?	None reported	NI
Questions relating to baseline and time-varying confounding		
1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding?	N/A	N/A
1.8. If Y/PY to 1.7: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?	N/A	N/A
Risk-of-bias judgement	For both OS and PFS The study is likely to have controlled for important known confounders that may not have been balanced across groups due to the design of the study. However, the report recognises that *Confounding due to unobserved factors is possible because patients were not randomised to treatment groups*	Moderate

Signalling questions		
Bias due to confounding	**Description**	**Response**
Optional: What is the predicted direction of bias due to confounding?	For both OS and PFS The study is sponsored by Novartis, the manufacturer of everolimus. This is not sufficient to assume the bias will be in favour of that drug and nothing specific suggests that was the case. The study purpose is to show that differences between the two drugs are not statistically significant	Unpredictable

Bias in selection of participants into the study		
2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention? If N/PN to 2.1: go to 2.4	Physicians randomly selected and abstracted data for up to five patient charts that met the prespecified inclusion criteria	**Y**
2.2. If **Y/PY** to 2.1: Were the post-intervention variables that influenced selection likely to be associated with intervention? 2.3 If **Y/PY** to 2.2: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome?	For inclusion in this study, patients were required to be at least 18 years of age, to have had an mRCC diagnosis, to have received a TKI (sunitinib, sorafenib or pazopanib) as first targeted therapy, and to have discontinued that therapy for medical reasons (e.g., drug intolerance, disease progression, and non-response without progression). In addition, patients were required to have subsequently initiated either everolimus or axitinib as second targeted therapy between 1 February 2012 and 1 January 2013	**PY** **PY**
2.4. Do start of follow-up and start of intervention coincide for most participants?	The patient's medical records were available for review from initiation of first targeted therapy until most recent follow-up or death	Y
2.5. If **Y/PY** to 2.2 and 2.3, or **N/PN** to 2.4: Were adjustment techniques used that are likely to correct for the presence of selection biases?		NI
Risk-of-bias judgement	For both OS and PFS as physician selection of patients	**Serious**
Optional: What is the predicted direction of bias due to selection of participants into the study?	For both OS and PFS	Unpredictable

Bias in classification of interventions		
3.1 Were intervention groups clearly defined?	Two intervention groups: everolimus or axitinib, although dose not specified	PY
3.2 Was the information used to define intervention groups recorded at the start of the intervention?	All data were taken from medical records	Y
3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome?	Intervention was specified as the second therapy	PN
Risk-of-bias judgement	For both OS and PFS	Low
Optional: What is the predicted direction of bias due to classification of interventions?	For both OS and PFS – not applicable because judged to be low risk of bias	N/A

Bias due to deviations from intended interventions		
If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2		
4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice?	The retrospective nature of this study means that participants were not allocated, but chosen on the basis of what they had already received in usual practice. However, number of discontinuations and reasons are not reported	PN
4.2. If **Y/PY** to 4.1: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome?	N/A	N/A
If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6		
4.3. Were important co-interventions balanced across intervention groups?	N/A	N/A
4.4. Was the intervention implemented successfully for most participants?	N/A	N/A
4.5. Did study participants adhere to the assigned intervention regimen?	N/A	N/A
4.6. If **N/PN** to 4.3, 4.4 or 4.5: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention?	N/A	N/A
Risk-of-bias judgement	For both OS and PFS	Low
Optional: What is the predicted direction of bias due to deviations from the intended interventions?	For both OS and PFS – not applicable as judged to be low risk of bias	N/A

Bias due to missing data		
5.1 Were outcome data available for all, or nearly all, participants?	The patient's medical records were available for review from initiation of first targeted therapy until most recent follow-up or death but it is unclear how many, if any, had no follow-up	NI
5.2 Were participants excluded due to missing data on intervention status?	For OS: patients without a recorded date of death at the time of medical records review were censored at the last recorded follow-up date	PN
	For PFS: patients without a recorded date of progression or death were censored at the last recorded follow-up date	PN
5.3 Were participants excluded due to missing data on other variables needed for the analysis?	No information reported in the publication	NI
5.4 If **PN/N** to 5.1, or **Y/PY** to 5.2 or 5.3: Are the proportion of participants and reasons for missing data similar across interventions?	N/A	N/A
5.5 If **PN/N** to 5.1, or **Y/PY** to 5.2 or 5.3: Is there evidence that results were robust to the presence of missing data?	N/A	N/A
Risk-of-bias judgement	For both OS and PFS	Low
Optional: What is the predicted direction of bias due to missing data?	For both OS and PFS – not applicable as judged to be low risk of bias	N/A

Bias in measurement of outcomes		
6.1 Could the outcome measure have been influenced by knowledge of the intervention received?	For OS: OS was defined as the time from the initiation of second targeted therapy to death from any cause; objective outcome measure. For PFS: PFS was defined as the time from the initiation of second targeted therapy to progression or death, whichever came first. Physicians assessed progression based on data available in the medical records, therefore risk of subjectivity in PFS assessment	N **Y**
6.2 Were outcome assessors aware of the intervention received by study participants?	No blinding in the study	**Y**
6.3 Were the methods of outcome assessment comparable across intervention groups?	'Patient data were anonymised and non-identifiable' for consent reasons. Different physicians involved in the study but otherwise no reason to suggest any differences between intervention groups for OS, but for PFS physicians did not have to use standard criteria	OS: PY **PFS: PN**
6.4 Were any systematic errors in measurement of the outcome related to intervention received?	None reported	NI
Risk-of-bias judgement	For OS: due to objectivity of outcome	Low
	For PFS: due to subjective nature of part of the outcome assessment	**Serious**
Optional: What is the predicted direction of bias due to measurement of outcomes?		Unpredictable

Bias in selection of the reported result		
Is the reported effect estimate likely to be selected, on the basis of the results, from . . .	For OS: discrete outcome of death from any cause	N
7.1. . . . multiple outcome *measurements* within the outcome domain?	For PFS: progression might have included radiographic evidence, physical exams, worsening performance status, worsening hypercalcemia, or growth of a subcutaneous or palpable mass, and report of cancer-related symptoms	**Y**
7.2. . . . multiple *analyses* of the intervention–outcome relationship?	Multivariable analyses presented with numerous individual variables applied separately and in a combined analysis	PN
7.3. . . . different *subgroups*?	Subgroup results presented alongside full cohort results	PN
Risk-of-bias judgement	For OS	Low
	For PFS: RECIST not used so could have been measured in various ways	**Serious**
Optional: What is the predicted direction of bias due to selection of the reported result?		Unpredictable

Overall bias		
Risk-of-bias judgement	For OS: due to physician selection of patients	**Serious**
	For PFS: due to physician selection of patients and physician decision on progression	**Serious**
	While the study controls for important known confounders and is considered low risk of bias for some domains, there are still risks of selection and confounding biases due to the retrospective design	
Optional: What is the overall predicted direction of bias for this outcome?		Unpredictable

Wong *et al.*[63]

Outcome

Overall survival and PFS – proposed benefits of the intervention.

Numerical result being assessed

Overall survival adjusted: HR 0.66, 95% CI 0.44 to 0.99; $p = 0.045$.

Progression-free survival adjusted: HR 0.76, 95% CI 0.55 to 1.37; $p = 0.931$.

Both adjusted using multivariate Cox PHs for: age, gender, race, whether or not metastasis was present at initial diagnosis, duration of mRCC, type of first targeted therapy, response to and duration of first targeted therapy, treatments received before first targeted therapy, comorbidities, number and sites of metastasis, sarcomatoid differentiation, non-clear-cell RCC, Karnofsky performance status, physician's practice setting and year of practice.

Responses underlined are potential markers for low risk of bias, and responses in **bold** are potential markers for a risk of bias. Where questions relate only to sign posts to other questions, no formatting is used.

Signalling questions		
Bias due to confounding	**Description**	**Response options**
1.1 Is there potential for confounding of the effect of intervention in this study?	Retrospective chart review of patients taking second-line TKI (sorafenib) or mTORi (everolimus)	**Y**
If N/PN to 1.1: the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered		
If **Y/PY** to 1.1: determine whether there is a need to assess time-varying confounding:		
1.2. Was the analysis based on splitting participants' follow-up time according to intervention received?		PN
If N/PN, answer questions relating to baseline confounding (1.4 to 1.6)		
If Y/PY, go to question 1.3		

Signalling questions		
Bias due to confounding	**Description**	**Response options**
1.3. Were intervention discontinuations or switches likely to be related to factors that are prognostic for the outcome?		N/A
If N/PN, answer questions relating to baseline confounding (1.4 to 1.6)		
If Y/PY, answer questions relating to both baseline and time-varying confounding (1.7 and 1.8)		
Questions relating to baseline confounding only		
1.4. Did the authors use an appropriate analysis method that controlled for all the important confounding domains?	OS and PFS HRs both adjusted for all the important confounding domains we identified except obesity, subsequent therapy, prior nephrectomy and age at diagnosis. Adjustments were made using multivariate Cox PHs	PY
1.5. If Y/PY to 1.4: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?	Mostly yes, although some not defined in enough detail to judge	PY
1.6. Did the authors control for any post-intervention variables that could have been affected by the intervention?		NI
Questions relating to baseline and time-varying confounding		
1.7. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding?	N/A	N/A
1.8. If Y/PY to 1.7: Were confounding domains that were controlled for measured validly and reliably by the variables available in this study?	N/A	N/A
Risk-of-bias judgement	There is some uncertainty about how some confounding domains were defined and built into the model, and there remains a possibility of unknown confounders biasing the effect due to the retrospective design	Moderate
Optional: What is the predicted direction of bias due to confounding?	Unknown	Unpredictable

Bias in selection of participants into the study		
2.1. Was selection of participants into the study (or into the analysis) based on participant characteristics observed after the start of intervention? If N/PN to 2.1: go to 2.4	Yes by definition since it was a chart review	**Y**
2.2. If **Y/PY** to 2.1: Were the post-intervention variables that influenced selection likely to be associated with intervention?	Patient outcomes may have influenced selection	**PY**
2.3 If **Y/PY** to 2.2: Were the post-intervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome?		**PY**
2.4. Do start of follow-up and start of intervention coincide for most participants?		Y

Bias in selection of participants into the study		
2.5. If **Y/PY** to 2.2 and 2.3, or **N/PN** to 2.4: Were adjustment techniques used that are likely to correct for the presence of selection biases?		NI
Risk-of-bias judgement	There are inherent selection biases involved in retrospective chart reviews	**Serious**
Optional: What is the predicted direction of bias due to selection of participants into the study?		Unpredictable

Bias in classification of interventions		
3.1 Were intervention groups clearly defined?	Yes, although not dose. Dose adjustments, discontinuations and subsequent therapy detailed	PY
3.2 Was the information used to define intervention groups recorded at the start of the intervention?	Yes and reviewed retrospectively	Y
3.3 Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome?	Unlikely – inclusion criteria defined clearly and participating physicians were blinded to the sponsor	PN
Risk-of-bias judgement		Low
Optional: What is the predicted direction of bias due to classification of interventions?	N/A – risk-of-bias judged to be low	N/A

Bias due to deviations from intended interventions		
If your aim for this study is to assess the effect of assignment to intervention, answer questions 4.1 and 4.2		
4.1. Were there deviations from the intended intervention beyond what would be expected in usual practice?	Retrospective so participants were not assigned but chosen on the basis of what they had already received in usual practice. Deviations from doses and additions to treatment sequences are detailed in table 3	PN
4.2. If **Y/PY** to 4.1: Were these deviations from intended intervention unbalanced between groups *and* likely to have affected the outcome?		N/A
If your aim for this study is to assess the effect of starting and adhering to intervention, answer questions 4.3 to 4.6		
4.3. Were important co-interventions balanced across intervention groups?		N/A
4.4. Was the intervention implemented successfully for most participants?		N/A
4.5. Did study participants adhere to the assigned intervention regimen?		N/A
4.6. If **N/PN** to 4.3, 4.4 or 4.5: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention?		N/A
Risk-of-bias judgement		Low
Optional: What is the predicted direction of bias due to deviations from the intended interventions?	N/A – risk-of-bias judged to be low	N/A

Bias due to missing data		
5.1 Were outcome data available for all, or nearly all, participants?	No flow diagram or information about data being missing	NI
5.2 Were participants excluded due to missing data on intervention status?		PN
5.3 Were participants excluded due to missing data on other variables needed for the analysis?	*Patients with missing value of any of the covariates were excluded from the analysis* *In all analyses, patients without observed death or progression events were censored at the date of last contact*	**PY**
5.4 If **PN/N** to 5.1, or **Y/PY** to 5.2 or 5.3: Are the proportion of participants and reasons for missing data similar across interventions?		NI
5.5 If **PN/N** to 5.1, or **Y/PY** to 5.2 or 5.3: Is there evidence that results were robust to the presence of missing data?		NI
Risk-of-bias judgement	Participants were excluded if there were any missing baseline data needed for the multivariate analyses. It is not clear how much data were missing for this reason, whether it varied across treatments, and whether or not results were robust	**Serious**
Optional: What is the predicted direction of bias due to missing data?		Unpredictable

Bias in measurement of outcomes		
6.1 Could the outcome measure have been influenced by knowledge of the intervention received?	OS: no PFS: 'progression in the present study was determined by treating physicians based on various diseases' monitoring methods and schedules used in real-world practice'	OS: N **PFS: PY**
6.2 Were outcome assessors aware of the intervention received by study participants?	Yes, although they were blinded to the sponsor	**Y**
6.3 Were the methods of outcome assessment comparable across intervention groups?	Not for OS Participating physicians may have had different practices for treatment sequence and PFS judgement	OS: PY **PFS: PN**
6.4 Were any systematic errors in measurement of the outcome related to intervention received?		NI
Risk-of-bias judgement	Some aspects are not well described but OS not likely to be biased by measurement	OS: low **PFS: serious**
Optional: What is the predicted direction of bias due to measurement of outcomes?	OS: N/A – rated as being at low risk of bias PFS: not able to judge – three groups and physicians assessing PFS were not aware of the sponsor	Unpredictable

Bias in selection of the reported result		
Is the reported effect estimate likely to be selected, on the basis of the results, from . . .		
7.1. . . . multiple outcome *measurements* within the outcome domain?	OS: No time point or type of measurement issues	OS: N
	PFS could have been measured in multiple ways	**PFS: PY**
7.2. . . . multiple *analyses* of the intervention–outcome relationship?	The paper states which data were excluded from analyses and which variables were included in the adjusted analyses	PN
7.3. . . . different *subgroups*?		PN
Risk-of-bias judgement		OS: low
		PFS: serious
Optional: What is the predicted direction of bias due to selection of the reported result?		Unpredictable

Overall bias		
Risk-of-bias judgement	While the study used some methods to control for biases, there were some that could not be avoided due to the retrospective design. There was also a risk of bias for missing data for the multivariate analyses, but we do not know how much was missing. Risks associated with outcome assessment and reporting are only present for PFS, not OS	**OS: serious** **PFS: serious**
Optional: What is the overall predicted direction of bias for this outcome?	While biases probably exist in several domains, the three-group design and blinding of participating physicians to the sponsor make it difficult to assess their direction	Unpredictable

Economic evaluation

Quality assessment of the included economic evaluations against the NICE reference case

		Comments								
Attribute	**Reference case**	**Study**								
		Paz-Ares *et al.*[83]	Petrou *et al.*[77]	Purmonen *et al.*[84]	Petrou[76]	Petrou *et al.*[78]	Mihajlovic *et al.*[79]	Hoyle *et al.*[81]	Casciano *et al.*[80]	Lopes *et al.*[85]
Decision problem	The scope developed by NICE	Yes	Yes	Yes	Yes	Yes	Yes	Partly	Partly	Partly
Comparator(s)	Alternative therapies routinely used in the NHS	Partly	Partly	Partly	Partly	Partly	Partly	Partly	Partly	N/A
Perspective costs	NHS and PSS	Partly. Non-UK but publicly funded health service	Partly. Non-UK but publicly funded health service	Partly. Non-UK but publicly funded health service	Partly. Non-UK but publicly funded health service	Partly. Non-UK but publicly funded health service	Partly. Non-UK but publicly funded health service	Yes	No. US payer perspective	No. US payer perspective
Perspective benefits	All health effects on individuals	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No. Only safety was incorporated in the analysis
Form of economic evaluation	Cost–utility analysis	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No. Budget impact analysis
Time horizon	Sufficient to capture differences in costs and outcomes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	6 years. No justification was provided for time horizon	Yes
Synthesis of evidence on outcomes	Systematic review	Unclear	Yes	Unclear	Yes	Unclear	Unclear	Yes	Not reported	No. Real-time drug utilisation data were used and AE rates were based a pivotal Phase III trial and drug prescribing data
Outcome measure	QALYs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	N/A

Attribute	Reference case	Comments								
		Study								
Health states for QALY	Described using a standardised and validated instrument	No. No description of health states. Utilities obtained directly from patients on sunitinib, before and after progression. Assumed to be the same for BSC group	Yes	Yes	Yes	Yes	Yes	Yes. EQ-5D	Not reported	N/A
Benefit valuation	TTO or standard gamble	No. EQ-VAS	Yes – UK TTO	Yes – TTO	Yes – UK TTO	Yes – UK TTO	Yes – UK TTO	Not reported	Not reported	N/A
Source of preference data for valuation of changes in HRQoL	Representative sample of the public	No. Directly elicited from patients	Yes – UK TTO	Yes – But does not specify which tariff is used	Yes – UK TTO	Yes – UK TTO	Yes – UK TTO	Not reported	Not reported	N/A
Discount rate	An annual rate of 3.5% on both costs and health effects	Yes	Yes	No – 5% discount rate was used	Yes	Yes	No. 1.5% for effects and 3% for costs	Yes	No. A rate of 3% for costs and outcomes was used	No discounting was reported which is standard practice in budget impact analyses
Equity	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	N/A
SA	PSA	Yes	Yes – base-case analysis is probabilistic	Yes – base-case analysis is probabilistic	Yes – base-case analysis is probabilistic	Yes – base-case analysis is probabilistic	Yes – base-case analysis is probabilistic	Yes	Yes	Not reported

N/A, not applicable; TTO, time trade-off.

Quality assessment of the included economic evaluations using the Philips checklist[87]

Dimension of quality	Comments					
	Study					
	Paz-Ares *et al.*[83]	Petrou *et al.*[77]	Purmonen *et al.*[84]	Petrou[76]	Petrou *et al.*[78]	Mihajlovic *et al.*[79]
Structure						
S1: Statement of decision problem/ objective	✓ Decision problem and objective clearly defined	✓ Decision problem and objective clearly stated	✓ Decision problem and objective are clearly stated	✓ Decision problem and objective are clearly stated	✓ Decision problem and objective are clearly stated	✓ Decision problem and objective are clearly stated
S2: Statement of scope/perspective	✓ Scope and perspective clearly stated	✓ Scope and perspective clearly stated	✓ Scope and perspective clearly stated	✓ Scope and perspective clearly stated	✓ Scope and perspective clearly stated	✓ Scope and perspective clearly stated
S3: Rationale for structure	✓ Based on a previously developed model from the USA and adapted for the Spanish healthcare service	✓ Model is clearly described alongside the disease pathways to show rationale. No other theories or evidence are described in considering the structure	✓ The model structure is clearly described and is appropriate for the decision problem. It is in line with other published economic models in this area	✓ The model structure is clearly described and is appropriate for the decision problem. It is in line with other published economic models in this area	✓ The model structure is clearly described and is appropriate for the decision problem. It is in line with other published economic models in this area	✓ The model structure is clearly described and is appropriate for the decision problem. It is in line with other published economic models in this area
S4: Structural assumptions	? Structural assumptions are stated but not all are clear. The cycle length of the Markov model is stated in three contradictory ways	✓ Structural assumptions are transparent, justified and reasonable	✓ Structural assumptions are transparent, justified and reasonable	✓ Structural assumptions are transparent, justified and reasonable	✓ Structural assumptions are transparent, justified and reasonable	✓ Structural assumptions are transparent, justified and reasonable
S5: Strategies/ comparators	✓ Comparators are clearly defined. Not all feasible options are evaluated as the scope was limited to one intervention compared with BSC	✓ Comparators are clearly defined. Not all feasible options are evaluated as the scope was limited to one intervention compared with BSC	✓ Comparators are clearly defined. Not all feasible options are evaluated as the scope was limited to one intervention compared with BSC	✓ Comparators are clearly defined. Not all feasible options are evaluated as the scope was limited to one intervention compared with BSC	✓ Comparators are clearly defined. Not all feasible options are evaluated as the scope was limited to one intervention compared with BSC	✓ Comparators are clearly defined. Not all feasible options are evaluated as the scope was limited to one intervention compared with BSC

Dimension of quality	Comments											
	Study											
S6: Model type	✓	Model type is appropriate	✓	Model type is appropriate	✓	Model type is appropriate	✓	Model type is appropriate	✓	Model type is appropriate	✓	Model type is appropriate
S7: Time horizon	✓	Based on an expert panel to allow almost 100% of patients to reach the state of death	✓	The time horizon is intended to allow all patients to reach the state of death	✓	Time horizon is a lifetime	✓	The time horizon is intended to allow all patients to reach the state of death	✓	The time horizon is intended to allow all patients to reach the state of death	✓	The time horizon is intended to allow all patients to reach the state of death
S8: Disease states/ pathways	✓	Health states reflect the underlying disease	✓	Health states reflect the underlying disease	✓	Health states reflect the underlying disease	✓	Health states reflect the underlying disease	✓	Health states reflect the underlying disease	✓	Health states reflect the underlying disease
S9: Cycle length	?	Cycle length is stated initially to fit with the 6-week treatment cycles but also later stated as 4-weekly and monthly	✓	The cycle length is defined monthly and is justified based on the low life expectancy	✓	Cycle length is clearly defined and reasonable	✓	Cycle length is clearly defined and reasonable	✓	Cycle length is clearly defined and reasonable	✓	Cycle length is clearly defined and reasonable
Data												
D1: Data identification	✗	Methods not fully described. Some unit costs are taken from studies without describing any search strategy or justification for including the evidence. Data quality was not reported as being assessed	✓	Clinical data were identified from a systematic review of literature	?	Methods for identifying data sources are not clearly detailed but references are given when necessary	✓	Clinical data were identified from a systematic review of literature	?	Methods for identifying data sources are not clearly detailed but references are given when necessary	?	Methods for identifying data sources are not clearly detailed but references are given when necessary
D2: Pre-model data analysis	?	No pre-model data analysis was reported	✓	Monthly transition probabilities were calculated from the trial data and these calculations are clearly described	✓	PFS and OS were estimated using a Weibull model. The methods are clearly described	✓	PFS and OS were used to estimate transition probabilities. Methods are clearly described	✓	PFS and OS were used to estimate transition probabilities. Methods are clearly described	✓	PFS and OS were used to estimate transition probabilities. Methods are clearly described

Dimension of quality	Comments											
	Study											
D2a: Baseline data	✗	Baseline data are taken from a retrospective database analysis. Details of calculations to derive transition probabilities are not reported	✓	All data were taken from one RCT of sorafenib versus BSC	✓	For the BSC strategy medical records from 39 patients from two Finnish University hospitals were used to represent survival and resource use. Methods for this are clearly described	✓	Baseline data for PFS, OS and utilities were taken from a Phase III RCT identified through systematic review	✓	Baseline PFS, OS and utilities were taken from a Phase III placebo controlled trial	✓	Baseline data for PFS and OS were taken from a RCT. Utilities were taken from published appraisals of other mRCC treatments
D2b: Treatment effects	✓	Treatment effects are taken from a single-arm Phase II trial	✗	Extrapolation of treatment effects is not clearly described	✓	Effects for sunitinib patients were taken from the pooled analysis of two Phase II single-arm trials. This was justified as no comparative studies were available	✓	Treatment effects were obtained from the RCT identified in the systematic review	✓	Treatment effects were taken from the same trial as the baseline data	✓	Treatment effects were taken from the same trial as the baseline data
D2c: Costs	✓	Costs included have been clearly stated along with assumptions made. Some costs were inflated using the consumer price index to estimate all costs in the same year	✓	Costs are justified. Discounting as per the NICE reference case	✓	Methods for estimating costs of BSC are referenced and detailed in full including methods for inflation and case-mix adjustment	✓	Costs are clearly described and sources given	✓	Costs are clearly described and sources given	✓	Costs are clearly described and sources given
D2d: Quality of life weights (utilities)	?	Methods of eliciting utilities are described but not justified. Utilities were only taken from patients in the intervention arm and assumed to apply to the equivalent health states in the BSC arm	✓	Quality-of-life data were found from literature and use the UK EQ-5D tariff	?	The EQ-5D is specified as the instrument used to derive utilities for patients in the intervention trial; however, the tariff used is not specified. Assumptions made to apply utilities to BSC are clearly stated	?	The EQ-5D was used to elicit utilities from patients in the trial used to inform the model. The tariff used is not stated	✓	The UK EQ-5D was used to elicit utilities from patients in the trial used to inform the model	?	The source of utility data is specified but the methods for eliciting are not given

Dimension of quality	Comments									
	Study									
D3: Data incorporation	✗	All data has been described with limited detail. Probability distributions are not reported for the PSA	✓	Data inputs and distributions have been described in sufficient detail	✓	Data are referenced and described. Distributions are detailed in full with justification for the type chosen and the parameter estimation	✓	Data inputs and distributions have been described in sufficient detail	✓	Data inputs and distributions have been described in sufficient detail	✓	Data inputs and distributions have been described in sufficient detail
D4: Assessment of uncertainty												
D4a: Methodological	✗	Assessment of methodological uncertainty has not been reported	✗	Assessment of methodological uncertainty has not been reported	✗	Assessment of methodological uncertainty has not been reported	✗	Assessment of methodological uncertainty has not been reported	✗	Assessment of methodological uncertainty has not been reported	✗	Assessment of methodological uncertainty has not been reported
D4b: Structural	✗	Assessment of structural uncertainty has not been reported	✗	Assessment of structural uncertainty has not been reported	✗	Assessment of structural uncertainty has not been reported	✗	Assessment of structural uncertainty has not been reported	✗	Assessment of structural uncertainty has not been reported	✗	Assessment of structural uncertainty has not been reported
D4c: Heterogeneity	✗	Heterogeneity has not been tested	✗	Heterogeneity has not been assessed	✓	Patients < 60 years and > 60 years were analysed separately in addition to the main analysis	✗	Heterogeneity has not been assessed	✗	Heterogeneity has not been tested	✗	Heterogeneity has not been tested

Dimension of quality	Comments									
	Study									
D4d: Parameter	?	A PSA was performed but the distributions around parameters have not been reported	✓	The model is probabilistic and 50,000 iterations were performed after discarding an initial set of 50,000 to allow for model stability	✓	The base-case model is probabilistic and OWSAs were also performed for the discount rate, time horizon and extending sunitinib treatment for another month. The values used for the discount rates and time horizon were not stated. A cost-effectiveness acceptability curve was also produced	✓	The base-case model is probabilistic and an initial 50,000 iterations of the model were discarded so that the model converges before results are analysed. EVPI (expected value of perfect information) was also performed	✓	The base-case model is probabilistic and an initial 50,000 iterations of the model were discarded so that the model converges before results are analysed. EVPI was also performed	✓	Parameter uncertainty was assessed through a PSA and various OWSAs using upper and lower 95% CIs or an arbitrary 20% change where these were not available
Consistency												
C1: Internal consistency	✓	The model was developed and tested in the USA and adapted for the Spanish health-care environment with a panel of local experts, including experts in economic evaluations	✗	It is not reported whether or not the mathematical logic in the model has been tested	✗	It is not reported whether or not the mathematical logic in the model has been tested	✗	It is not reported whether or not the mathematical logic in the model has been tested	✗	It is not reported whether or not the mathematical logic in the model has been tested	✗	It is not reported whether or not the mathematical logic in the model has been tested
C2: External consistency	✓	The results are compared with other studies and are similar	✗	It is not reported whether or not external consistency has been tested	✓	The time required for the model to reach completion was compared with empirical studies and appeared to be consistent	✗	It is not reported whether or not external consistency has been tested	✗	It is not reported whether or not external consistency has been tested	✗	It is not reported whether or not external consistency has been tested

Dimension of quality	Comments					
	Study					
	Hoyle 2010[81]		Casciano 2011[80]		Lopes 2012[85]	
Structure						
S1: Statement of decision problem/ objective	✓	Stated clearly	✓	Stated clearly	✓	Stated clearly
S2: Statement of scope/perspective	✓	Stated clearly, UK NHS/ PSS perspective	✓	Stated clearly, US payer perspective	✓	Stated clearly, US payer perspective
S3: Rationale for structure	✓	Stated clearly	✗	Not stated	✗	Not stated/not applicable
S4: Structural assumptions	✓	Stated clearly	✓	Stated clearly	✗	Not stated/not applicable
S5: Strategies/ comparators	?	Did not include the full range of comparators but considered sorafenib and BSC	?	Did not include the full range of comparators but considered everolimus compared with sorafenib after failure of treatment with sunitinib	?	Did not include the full range of comparators but considered introduction of everolimus as second/ third line treatment option
S6: Model type	✓	Markov model	✓	Markov model	✓	Cross-sectional budget impact model
S7: Time horizon	✓	Life-time horizon	✓	6 years, no justification was provided for this time horizon	✓	Costs were estimated for the periods of April 2008 to March 2009, and October 2009 to September 2010 to reflect the periods before and after expected uptake of everolimus
S8: Disease states/ pathways	✓	PFS, PD and death	✓	Stable disease with no AEs, stable disease with AEs, disease progression and death. No clarification was given with regards to having separate stable disease states (with or without AEs)	✗	Not stated
S9: Cycle length	✓	6 weeks, no rationale was provided for this duration	✓	8 weeks, to reflect the time period of assessment in the trial	✗	Not stated/not applicable
Data						
D1: Data identification	✓	Reported	✓	Partly	✓	Stated clearly
D2: Pre-model data analysis	✓	Reported	✗	Not reported	✓	Reported
D2a: Baseline data	✓	Reported	✗	Not reported	✗	Not reported
D2b: Treatment effects	✓	Relative treatment effects were reported for PFS and OS	✗	The outcomes used were not explicitly reported	✗	Not reported/not applicable
D2c: Costs	✓	Details of how costs were calculated in the model were reported	✓	Details of how costs were calculated in the model were reported	✓	Details of how costs were calculated in the model were reported

Dimension of quality	Comments					
	Study					
D2d: Quality of life weights (utilities)	?	The authors reported that utility values were derived from a Phase II trial of sunitinib as reported in Motzer *et al.* 2006.[92] However, they did not have access to the EQ-5D data used to estimate the health-state utility values	?	Quality of life weights were obtained from published literature. No details were provided on how the papers were chosen	✗	Not reported/not applicable
D3: Data incorporation	✓	Reported	✗	It is not possible to validate how the data were incorporated due to lack of reporting	✗	It is not possible to validate the incorporation of data due to a lack of reporting
D4: Assessment of uncertainty						
D4a: Methodological	✗	Not reported	✗	Not reported	✗	Not reported
D4b: Structural	✗	Not reported	✓	Structural uncertainty was explored by changing assumptions surrounding RDI and mortality rate after progression as part of the deterministic SA	✗	Not reported/not applicable
D4c: Heterogeneity	✗	Not reported	✗	Not reported	✗	Not reported
D4d: Parameter	✓	The effect of parameter uncertainty on cost-effectiveness was explored through one-way and PSAs	✓	Parametric uncertainty was explored through deterministic SAs and a PSA around the base case	✓	Scenario analyses were carried out
Consistency						
C1: Internal consistency	✓	The authors reported that the cost-effectiveness model was verified	✗	Measures taken to ensure internal consistency was not reported	✗	Not reported
C2: External consistency	✗	The results of the analysis have not been compared with the results of the trials informing them or to published cost-effectiveness papers	✗	The results of the analysis have not been compared with the results of the trials informing them or to published cost-effectiveness papers	✗	Not reported

Appendix 10 Table of excluded studies with rationale

Clinical literature

Randomised controlled trial search

Paper excluded	Full reference details	Reason for exclusion
Albiges 2011	Albiges L, Antoun S, Martin L, Merad M, Loriot Y, Baracos VE, *et al.* Effect of everolimus therapy on skeletal muscle wasting in patients with metastatic renal cell carcinoma (mRCC): results from a placebo-controlled study. Genitourinary Cancers Symposium; 2011. *J Clin Oncol* 2011;**29** (Suppl.7):319	Ineligible data
Ambring 2011	Ambring AE, Stierner UK, Oden AS, Bjorholt IN. Sorafenib and sunitinib in renal cell cancer: a study based on register data. *J Clin Oncol* 2011;**29**(Suppl. 15):4600	Ineligible data
National Horizon Scanning Centre 2006	NHSC. *Sorafenib Tosylate (Nexavar) for Advanced Renal Cell Carcinoma: Horizon Scanning Review*. Birmingham: National Horizon Scanning Centre; 2006:**5**	Ineligible study design
National Horizon Scanning Centre 2008	NHSC. *Everolimus for Advanced and/or Metastatic Renal Cell Carcinoma – Second Line*. Birmingham: National Horizon Scanning Centre; 2008	Ineligible study design
Anonymous 2008	Anonymous. Immunosuppresant everolimus improves progression-free survival in advanced kidney cancer patients. *Oncology* 2008;**22**:841	Ineligible study design
Anonymous 2008	Anonymous. Renal cell carcinoma: everolimus prolongs progression-free survival. *Arzneimitteltherapie*. 2008;**26**:307–8	Ineligible study design
Anonymous 2008	Anonymous. The multikinase inhibitor sorafenib prolongs survival in kidney and liver cancer. *Onkologie* 2008;**31**:205	Ineligible study design
National Horizon Scanning Centre 2010	National Horizon Scanning Centre. *Axitinib for Advanced and/or Metastatic Renal Cell Carcinoma - Second Line*. Birmingham: National Horizon Scanning Centre; 2010	Ineligible study design
Anonymous 2010	Anonymous. Advanced renal cell cancer: significance of the sequence therapy for optimal treatment success. *Onkologie* 2010;**33**:270–1	Ineligible study design
NICE 2011	Everolimus for the second-line treatment of advanced renal cell carcinoma (Structured abstract). *Health Technology Assessment Database*. 2016; Issue 4. URL: http://onlinelibrary.wiley.com/o/cochrane/clhta/articles/HTA-32011000520/frame.html	Ineligible data
Antoun 2010	Antoun S, Birdsell L, Sawyer MB, Venner P, Escudier B, Baracos VE. Association of skeletal muscle wasting with treatment with sorafenib in patients with advanced renal cell carcinoma: results from a placebo-controlled study. *J Clin Oncol* 2010;**28**:1054–60	Ineligible data
Antoun 2011	Antoun S, Albiges L, Martin L, Merad-Taoufik M, Baracos VE, Escudier B. Effect of everolimus an anti mtor therapy, on skeletal muscle wasting in patients with metastatic renal cell carcinoma (MRCC). *Supportive Care in Cancer* 2011;**19**(Suppl. 1):S161	Ineligible data
Bellmunt 2009	Bellmunt J. Future developments in renal cell carcinoma. *Ann Oncol* 2009;**20**(Suppl. 1):i13-i7	Ineligible study design

Paper excluded	Full reference details	Reason for exclusion
Blute 2006	Blute ML. Sunitinib in patients with metastatic renal cell carcinoma. Motzer RJ, Rini BI, Bukowski RM, Curti BD, George DJ, Hudes GR, Redman BG, Margolin KA, Merchan JR, Wilding G, Ginsberg MS, Bacik J, Kim ST, Baum CM, Michaelson MD, Department of Medicine, Memorial Sloan-Kettering Cancer Centre, New York, NY. *Urologic Oncology: Seminars and Original Investigations* 2006;**24**:553–4	Ineligible study design
Bracarda 2012	Bracarda S, Hutson TE, Porta C, Figlin RA, Calvo E, Grunwald V, *et al.* Everolimus in metastatic renal cell carcinoma patients intolerant to previous VEGFr-TKI therapy: a RECORD-1 subgroup analysis. *Br J Cancer* 2012;**106**:1475–80	Ineligible population
Bukowski 2007	Bukowski R, Cella D, Gondek K, Escudier B. Effects of sorafenib on symptoms and quality of life: results from a large randomised placebo-controlled study in renal cancer. *Am J Clin Oncol* 2007;**30**:220–7	Ineligible population
Bukowski 2009	Bukowski R, Eisen T, Stadler T, Szczylic C, Oudard S, Siebels M, *et al.* Efficacy and safety of sorafenib in patients with advanced clear-cell renal-cell carcinoma (RCC) with bone metastases: results from the phase III target study. *Eur J Cancer* 2009;**7**:432	Ineligible data
Cella 2011	Cella D, Escudier B, Rini BI, Chen C, Bhattacharyya H, Tarazi JC, *et al.* Patient-reported outcomes (PROs) in a phase III AXIS trial of axitinib versus sorafenib as second-line therapy for metastatic renal cell carcinoma (mRCC). *J Clin Oncol* 2011;**29**(Suppl. 1)	Ineligible data
Cella 2011	Cella D, Escudier B, Rini B, Chen C, Bhattacharyya H, Tarazi J, *et al. Time to Deterioration (TTD) in Patient-reported Outcomes in Phase 3 Axis Trial of Axitinib vs Sorafenib as Second-line Therapy for Metastatic Renal Cell Carcinoma (mRCC). Eur J Cancer* European Multidisciplinary Cancer Congress, Stockholm, Sweden, 2011	Ineligible data
Choueiri 2010	Choueiri TK, Schutz FA, Je Y, Rosenberg JE, Bellmunt J. Risk of arterial thromboembolic events with sunitinib and sorafenib: a systematic review and meta-analysis of clinical trials. *J Clin Oncol* 2010;**28**:2280–5	Ineligible study design
Chu 2009	Chu D, Lacouture ME, Weiner E, Wu S. Risk of hand-foot skin reaction with the multitargeted kinase inhibitor sunitinib in patients with renal cell and non-renal cell carcinoma: a meta-analysis (structured abstract). *Clin Genitourin Cancer* 2009;**7**:11–19	Ineligible population
Coon 2010	Thompson Coon J, Hoyle M, Green C, Liu Z, Welch K, Moxham T. Bevacizumab, sorafenib tosylate, sunitinib and temsirolimus for renal cell carcinoma: a systematic review and economic evaluation. *Health Technol Assess* 2010;**14**(2).	Ineligible population
Coppin 2008	Coppin C, Le L, Porzsolt F, Wilt T. Targeted therapy for advanced renal cell carcinoma. *Cochrane Database Syst Rev* 2008;**16**:CD006017	Ineligible study design
Dhanda 2006	Dhanda R, Gondek K, Song J, Cella D, Bukowski RM, Escudier B. A comparison of quality of life and symptoms in kidney cancer patients receiving sorafenib versus placebo. *J Clin Oncol* 2006;**24**:4534	Ineligible data
Di Lorenzo 2011	Di Lorenzo G, Casciano R, Malangone E, Buonerba C, Sherman S, Willet J, *et al.* An adjusted indirect comparison of everolimus and sorafenib therapy in sunitinib-refractory metastatic renal cell carcinoma patients using repeated matched samples. [Erratum appears in *Expert Opin Pharmacother* 2013;**14**:2003.] [Erratum appears in *Expert Opin Pharmacother* 2011;**12**:2143.] *Expert Opin Pharmacother* 2011;**12**:1491–7	Ineligible data
Diaz 2015	Diaz J, Sternberg CN, Mehmud F, Delea TE, Latimer N, Bartlett-Pandite AN, *et al.* Crossover in oncology clinical trials. *J Clin Oncol* 2015;**33**(Suppl. 1)	Ineligible study design

Paper excluded	Full reference details	Reason for exclusion
Eichelberg 2012	Eichelberg C, Fischer Von Weikersthal L, Goebell P, Lerchenmuller C, Zimmermann U, Freier W, *et al.* Phase III randomised sequential open-label study to evaluate efficacy and safety of sorafenib (SO) followed by sunitinib (SU) vs. sunitinib followed by sorafenib in patients with advanced/meta-static renal cell carcinoma (mRCC) without prior systemic therapy (SWITCH Study) – safety interim analysis results. *Urologe – Ausgabe A* 2012;**51**:35	Ineligible data
Eisen 2006	Eisen T, Bukowski RM, Staehler M, Szczylik C, Oudard S, Stadler WM, *et al.* Randomised phase III trial of sorafenib in advanced renal cell carcinoma (RCC): impact of crossover on survival. *J Clin Oncol* 2006;**24**:4524	Ineligible population
Eisen 2008	Eisen T, Oudard S, Szczylik C, Gravis G, Heinzer H, Middleton R, *et al.* Sorafenib for older patients with renal cell carcinoma: subset analysis from a randomised trial. *J Nat Cancer Institute* 2008;**100**:1454–63	Ineligible population
Escudier 2007	Escudier B, Eisen T, Stadler WM, Szczylik C, Oudard S, Siebels M, *et al.* Sorafenib in advanced clear-cell renal-cell carcinoma. *N Engl J Med* 2007;**356**:125–34	Ineligible population
Escudier 2007	Escudier B. Sorafenib in kidney cancer. *Ann Oncol* 2007;**18**(Suppl. 9)	Ineligible population
Escudier 2009	Escudier B, Eisen T, Stadler WM, Szczylik C, Oudard S, Staehler M, *et al.* Sorafenib for treatment of renal cell carcinoma: final efficacy and safety results of the phase III treatment approaches in renal cancer global evaluation trial. *J Clin Oncol* 2009;**27**:3312–18	Ineligible population
Escudier 2012	Escudier B, Rini BI, Hutson TE, Gore M, Oudard S, Tarazi J, *et al.* Updated results of the phase 3 AXIS trial: Axitinib vs sorafenib as second-line therapy for metastatic renal cell carcinoma (mRCC). *Eur Urol* 2012;**11**:e81-ea	Ineligible data
Goebell 2014	Goebell PJ, Vervenne W, Santis M, Weikersthal LF, Lerchenmuller CA, Zimmermann U, *et al.* Subgroup analyses of a randomised sequential open-label study (SWITCH) to evaluate efficacy and safety of sorafenib (SO)/sunitinib (SU) versus SU/SO in the treatment of metastatic renal cell cancer (mRCC). *J Clin Oncol* 2014;**32**(Suppl. 1)	Ineligible population
Gschwend 2010	Gschwend J, Bukowski R, Eisen T, Stadler W, Szczylik C, Oudard S, *et al.* Efficacy and safety of sorafenib in patients with advanced clear-cell renal-cell carcinoma (RCC) with bone metastases: results from the phase III TARGET study. *Onkologie* 2010;**33**:130–1	Ineligible data
Hsieh 2015	Hsieh J, Chen D, Wang P, Chen Y, Redzematovic A, Marker M, *et al.* Identification of efficacy biomarkers in a large metastatic renal cell carcinoma (mRCC) cohort through next generation sequencing (NGS): results from RECORD-3. *J Clin Oncol* 2015;**33**(Suppl. 1)	Ineligible data
Hutson 2010	Hutson TE, Bellmunt J, Porta C, Szczylik C, Staehler M, Nadel A, *et al.* Long-term safety of sorafenib in advanced renal cell carcinoma: follow-up of patients from phase III TARGET. *Eur J Cancer* 2010;**46**:2432–40	Ineligible population
Hutson 2011	Hutson TE, Bracarda S, Escudier B, Porta C, Figlin RA, Calvo E, *et al.* Phase III, randomised, placebo-controlled study of everolimus in patients with metastatic renal cell carcinoma (mRCC): subgroup analysis of patients intolerant of prior vascular endothelial growth factor receptor-tyrosine kinase inhibitor (VEGFr-TKI) therapy. *J Clin Oncol* 2011;**29**(Suppl. 1)	Ineligible population
Hutson 2014	Hutson T, Bukowski R, Rini B, Gore M, Larkin J, Figlin R, *et al.* Efficacy and safety of sunitinib in elderly patients with metastatic renal cell carcinoma. *Br J Cancer* 2014;**110**:1125–32	Ineligible intervention
Ibrahim 2013	Ibrahim EM, Kazkaz GA, Abouelkhair KM, Bayer AM, Elmasri OA. Sunitinib adverse events in metastatic renal cell carcinoma: a meta-analysis. *Int J Clin Oncol* 2013;**18**:1060–9	Ineligible study design

Paper excluded	Full reference details	Reason for exclusion
Je 2009	Je Y, Schutz FA, Choueiri TK. Risk of bleeding with vascular endothelial growth factor receptor tyrosine-kinase inhibitors sunitinib and sorafenib: a systematic review and meta-analysis of clinical trials. *Lancet Oncol* 2009;**10**:967–74	Ineligible population
Kenney 2012	Kenney PA, Wood CG. Re: Comparative effectiveness of axitinib versus sorafenib in advanced renal cell carcinoma (AXIS): a randomised phase 3 trial. *Eur Urol* 2012;**62**:182–3	Ineligible study design
Kim 2009	Kim A, Balis FM, Widemann BC. Sorafenib and sunitinib. *Oncologist* 2009;**14**:800–5	Ineligible study design
Knox 2010	Knox JJ, Kay AC, Schiff E, Hollaender N, Rouyrre N, Ravaud A, *et al.* First-line everolimus followed by second-line sunitinib versus the opposite treatment sequence in patients with metastatic renal cell carcinoma (mRCC). *J Clin Oncol* 2010;**28**(Suppl. 15):39	Ineligible data
Leung 2011	Leung HW, Chan AL. Multikinase inhibitors in metastatic renal cell carcinoma: indirect comparison meta-analysis. *Clin Ther* 2011;**33**:708–16	Ineligible study design
Mills 2009	Mills EJ, Rachlis B, O'Regan C, Thabane L, Perri D. Metastatic renal cell cancer treatments: an indirect comparison meta-analysis. *BMC Cancer* 2009;**9**:34	Ineligible study design
Motzer 2008	Motzer RJ, Escudier B, Oudard S. RAD001 vs placebo in patients with metastatic renal cell carcinoma after progression on VEGFr-TKI therapy: results from a randomised, double-blind, multicenter phase-III study. *J Clin Oncol* 2008;**26**(Suppl.)	Ineligible data
Motzer 2013	Motzer RJ, Barrios CH, Kim TM, Falcon S, Cosgriff T, Harker WG, *et al.* Record-3: phase II randomised trial comparing sequential first-line everolimus (EVE) and second-line sunitinib (SUN) versus first-line SUN and second-line EVE in patients with metastatic renal cell carcinoma (mRCC). *J Clin Oncol* 2013;**31**(15 Suppl. 1)	Ineligible population
Motzer 2014	Motzer RJ, Barrios CH, Kim TM, *et al.* Phase II randomised trial comparing sequential first-line everolimus and second-line sunitinib versus first-line sunitinib and second-line everolimus in patients with metastatic renal cell carcinoma. *J Clin Oncol* 2014;**32**:2765–72	Ineligible data
Nachtnebel 2009	Nachtnebel A. Everolimus (Afinitor) for advanced/metastatic kidney cancer (structured abstract). *Health Technology Assessment Database* 2016; Issue 4. URL: http://eprints.hta.lbg.ac.at/857/	Ineligible data
NIHR 2014	NIHR Horizon Scanning Centre. Nivolumab for advanced or metastatic clear-cell renal cell carcinoma? Second or third line. *Health Technology Assessment Database* 2016; Issue 4. URL: http://onlinelibrary.wiley.com/o/cochrane/clhta/articles/HTA-32015000122/frame.html	Ineligible data
Oudard 2009	Oudard S, Eisen T, Szczylik C, Negrier S, Chevreau C, Cihon F, *et al.* Efficacy and safety of sorafenib in patients with advanced clear-cell renal cell carcinoma (RCC) with diabetes: results from the phase III TARGET study. *J Clin Oncol* 2009;**27**(Suppl. 1):e16099	Ineligible population
Oudard 2011	Oudard S, Escudier B, Hutson T, Porta C, Bracarda S, Grunwald V, *et al.* Everolimus in patients with metastatic renal cell carcinoma: subgroup analysis of patients with a reduction in tumour burden enrolled in a randomised, placebo-controlled, phase III trial. *Eur Urol* 2011;**10**:229	Ineligible data
Oudard 2012	Oudard S, Escudier B, Thompson J, Grunwald V, Conte P, Bracarda S, *et al.* Biomarkers of everolimus efficacy in patients with Metastatic Renal Cell Carcinoma (MRCC): analysis of the phase III RECORD-1 trial. *Ann Oncol* 2012;**23**:ix278–ix9	Ineligible data
Oudard 2013	Oudard S, Escudier B, Thompson J, Grunwald V, Masini C, Bracarda S, *et al.* Relationship between biomarkers and everolimus efficacy in the phase III RECORD-1 trial of patients with metastatic renal cell carcinoma (mRCC). *J Clin Oncol* 2013;**31**:352	Ineligible data

Paper excluded	Full reference details	Reason for exclusion
Peña 2010	Peña C, Lathia C, Shan M, Escudier B, Bukowski RM. Biomarkers predicting outcome in patients with advanced renal cell carcinoma: results from sorafenib phase III treatment approaches in renal cancer global evaluation trial. *Clin Cancer Res* 2010;**16**:4853–63	Ineligible data
Poggiani 2012	Poggiani C, Hintringer K. Axitinib for 2nd-line metastatic renal cell carcinoma. *Health Technology Assessment Database* 2016; Issue 4. URL: http://onlinelibrary.wiley.com/o/cochrane/clhta/articles/HTA-32012000309/frame.html	Ineligible data
Porta 2011	Porta C, Escudier B, Hutson T, Figlin R, Calvo E, Grunwald V, *et al.* Analysis of the relationship between Karnofsky performance status (KPS) and tumour response in the RECORD-1 phase III trial of everolimus in patients with advanced renal cell carcinoma (RCC). *J Clin Oncol* 2011;**29**:4610	Ineligible data
Porta 2012	Porta C, Calvo E, Climent MA, Vaishampayan U, Osanto S, Ravaud A, *et al.* Efficacy and safety of everolimus in elderly patients with metastatic renal cell carcinoma: an exploratory analysis of the outcomes of elderly patients in the RECORD-1 Trial. *Eur Urol* 2012;**61**:826–33	Ineligible population
Porta 2012	Porta C, Escudier B, Hutson T, Figlin R, Calvo E, Grunwald V, *et al.* Relationship between karnofsky performance status (KPS) and tumour response: analysis of the RECORD-1 phase 3 trial of everolimus in patients with advanced renal cell carcinoma (RCC). *BJU Int* 2012;**109**:9–10	Ineligible data
Qu 2012	Qu AQ, Cheng SC, Atkins M, Signoretti S, Choueiri TK. Carbonic anhydrase IX (CAIX) as a potential biomarker of efficacy in metastatic clear-cell renal cell carcinoma (mccRCC) in patients (pts) receiving sorafenib: analysis of a randomised controlled trial (TARGET). *J Clin Oncol* 2012;**30**:352	Ineligible data
Rexer 2012	Rexer H. First-line therapy of advanced or metastasized renal cell cancer: open randomised phase III sequence study to examine the effectiveness and tolerance of sorafenib followed by pazopanib versus pazopanib followed by sorafenib in the first-line treatment of patients with advanced or metastasized renal cell cancer (SWITCH-2 – AN 33/11). *Der Urologe Ausg A* 2012;**51**:724–6	Ineligible intervention
Rexer 2014	Rexer H, Auo. First-line therapy of advanced or metastasized renal cell carcinoma: phase III, open, randomised sequence study to examine efficacy and tolerance of sorafenib followed by pazopanib versus pazopanib followed by sorafenib in the first-line treatment of patients with advanced or metastasized renal cell carcinoma (SWITCH-2 – AN 33/11). *Der Urologe Ausg A* 2014;**53**:735–8	Ineligible intervention
Richards 2011	Richards CJ, Je Y, Schutz FA, Heng DY, Dallabrida SM, Moslehi JJ, *et al.* Incidence and risk of congestive heart failure in patients with renal and nonrenal cell carcinoma treated with sunitinib. *J Clin Oncol* 2011;**29**:3450–6	Ineligible study design
Rosenbaum 2008	Rosenbaum SE, Wu S, Newman MA, West DP, Kuzel T, Lacouture ME. Dermatological reactions to the multitargeted tyrosine kinase inhibitor sunitinib. *Supportive Care in Cancer* 2008;**16**:557–66	Ineligible study design
Russo 2006	Russo P. Phase II placebo-controlled randomised discontinuation trial of sorafenib in patients with metastatic renal cell carcinoma. Ratain MJ, Eisen T. Stadler WM, Flaherty KT, Kaye SB, Rosner GL, Gore M, Desai A, Patnaik A, Xiong HQ, Rowinsky E, Abbruzzese JL, Xia C, Simantov R, Schwartz B. O'Dwyer PJ, University of Chicago, Chicago, IL. *Urologic Oncology: Seminars and Original Investigations* 2006;**24**:[560 *p.*]	Ineligible study design
Schmidinger 2011	Schmidinger M, Vogl UM, Bojic M, Lamm W, Heinzl H, Haitel A, *et al.* Hypothyroidism in patients with renal cell carcinoma. *Cancer* 2011;**117**:534–44	Ineligible population
Sivendran 2012	Sivendran S, Liu Z, Portas LJ, Yu M, Hahn N, Sonpavde G, *et al.* Treatment-related mortality with vascular endothelial growth factor receptor tyrosine kinase inhibitor therapy in patients with advanced solid tumours: a meta-analysis. *Cancer Treat Rev* 2012;**38**:919–25	Ineligible study design

Paper excluded	Full reference details	Reason for exclusion
Stein 2011	Stein AM, Carter A, Hollaender N, Motzer RJ, Sarr C. Quantifying the effect of everolimus on both tumour growth and new metastases in metastatic renal cell carcinoma (RCC): a dynamic tumour model of the RECORD-1 phase III trial. *J Clin Oncol* 2011;**29**(Suppl. 1)	Ineligible data
Stein 2012	Stein A, Wang W, Carter AA, Chiparus O, Hollaender N, Kim H, *et al.* Dynamic tumour modelling of the dose–response relationship for everolimus in metastatic renal cell carcinoma using data from the phase 3 RECORD-1 trial. *BMC Cancer* 2012;**12**:311	Ineligible data
Stein 2013	Stein A, Bellmunt J, Escudier B, Kim D, Stergiopoulos SG, Mietlowski W, *et al.* Survival prediction in everolimus-treated patients with metastatic renal cell carcinoma incorporating tumour burden response in the RECORD-1 trial. *Eur Urol* 2013;**64**:994–1002	Ineligible data
Stenner 2012	Stenner F, Chastonay R, Liewen H, Haile SR, Cathomas R, Rothermundt C, *et al.* A pooled analysis of sequential therapies with sorafenib and sunitinib in metastatic renal cell carcinoma. *Oncology* 2012;**82**:333–40	Ineligible study design
Thiam 2010	Thiam R, Cuenod CA, Fournier L, Lamuraglia M, Medioni J, Barascout B, *et al.* Determination of a new RECIST threshold using everolimus treatment in metastatic renal cell carcinoma: evaluation from the RECORD-1 study. *Ann Oncol* 2010;**21**:viii74–viii5	Ineligible population
Uemura 2012	Uemura H, Ou YC, Lim HY, Tomita Y, Ueda T, Menon H, *et al.* Phase III axis trial of axitinib versus sorafenib in patients with metastatic renal cell carcinoma: Asian subgroup analysis. *Ann Oncol* 2012;**23**:xi6	Ineligible population
Vickers 2010	Vickers MM, Choueiri TK, Rogers M, Percy A, Finch D, Zama I, *et al.* Clinical outcome in metastatic renal cell carcinoma patients after failure of initial vascular endothelial growth factor-targeted therapy. *Urology* 2010;**76**:430–4	Ineligible intervention
Voss 2014	Voss MH, Chen D, Marker M, Hsieh J, Knox JJ, Anak O, *et al.* A composite score of 5 circulating biomarkers predicts benefit from everolimus: results from 442 patients (pts) randomised on RECORD-3. *BJU Int* 2014;**114**:17	Ineligible data
Voss 2014	Voss MH, Chen D, Marker M, Hamilton M, Kalfoglou C, Hsieh J, *et al.* Identification and validation of predictive biomarkers (BM) for everolimus (EVE) in metastatic renal cell carcinoma: analysis of 442 patients on RECORD-3. *J Clin Oncol* 2014;**32**(Suppl. 15):4531	Ineligible data
Voss 2016	Voss MH, Chen D, Marker M, Hakimi AA, Lee CH, Hsieh JJ, *et al.* Circulating biomarkers and outcome from a randomised phase II trial of sunitinib vs everolimus for patients with metastatic renal cell carcinoma. *Br J Cancer* 2016;**114**:642–9	Ineligible data
Wu 2008	Wu S, Chen JJ, Kudelka A, Lu J, Zhu X. Incidence and risk of hypertension with sorafenib in patients with cancer: a systematic review and meta-analysis. *Lancet Oncol* 2008;**9**:117–23	Ineligible study design
Yousaf 2013	Yousaf N, Larkin J. Axitinib in advanced renal-cell carcinoma. *Lancet Oncol* 2013;**14**:1245–6	Ineligible study design

Observational study search

Paper excluded	Full reference details	Reason for exclusion
Albiges 2012	Albiges L, Riet F, Massard C, Le Moulec S, Loriot Y, Levy A, *et al.* Second line treatment in metastatic papillary renal cell carcinoma: retrospective analysis of a 48 patients cohort. *Ann Oncol* 2012;**23**:ix277	Ineligible intervention
Albiges 2015	Albiges L, Choueiri T, Escudier B, Galsky M, George D, Hofmann F, *et al.* A systematic review of sequencing and combinations of systemic therapy in metastatic renal cancer. *Eur Urol* 2015;**67**:100–10	Ineligible study design
Alimohamed 2014	Alimohamed N, Lee JL, Srinivas S, Bjarnason GA, Knox JJ, Mackenzie MJ, *et al.* A population-based overview of sequences of targeted therapy in metastatic renal cell carcinoma. *Clin Genitourin Cancer* 2014;**12**: e127–31	Ineligible data
Ambring 2011	Ambring AE, Stierner UK, Oden AS, Bjorholt IN. Sorafenib and sunitinib in renal cell cancer: a study based on register data. *J Clin Oncol* 2011;**29**(Suppl. 15):4600	Ineligible data
Antonelli 2011	Antonelli A, Daja J, Ferrari V, Arrighi N, Cunico SC, Simeone C. Sequential target therapy for metastatic renal cell carcinoma: comparison of sunitinib+sorafenib vs. sorafenib+sunitinib. *Anticancer Res* 2011;**31**:1922–3	Ineligible data
Autier 2008	Autier J, Escudier B, Wechsler J, Spatz A, Robert C. Prospective study of the cutaneous adverse effects of sorafenib, a novel multikinase inhibitor. *Arch Derm* 2008;**144**:886–92	Ineligible population
Biondani 2014	Biondani P, Verzoni E, Torri V, Porcu L, Grassi P, Testa I, *et al.* Sequential Tyrosine Kinase Inhibitors (TKIs) in metastatic renal cell carcinoma: results from a large cohort of patients. *Anticancer Res* 2014;**34**:2395–8	Ineligible data
Buchler 2012	Buchler T, Klapka R, Melichar B, Brabec P, Dusek L, Vyzula R, *et al.* Sunitinib followed by sorafenib or vice versa for metastatic renal cell carcinoma – data from the Czech registry. *Ann Oncol* 2012;**23**:395–401	Ineligible data
Buchler 2012	Buchler T, Pavlik T, Bortlicek Z, Poprach A, Vyzula R, Abrahamova J, *et al.* Objective response and time to progression on sequential treatment with sunitinib and sorafenib in metastatic renal cell carcinoma. *Med Oncol* 2012;**29**:3321–4	Ineligible data
Busch 2011	Busch J, Seidel C, Weikert S, Wolff I, Kempkensteffen C, Weinkauf L, *et al.* Intrinsic resistance to tyrosine kinase inhibitors is associated with poor clinical outcome in metastatic renal cell carcinoma. *BMC Cancer* 2011;**11**:295	Ineligible population
Busch 2011	Busch J, Seidel C, Kempkensteffen C, Johannsen M, Wolff I, Hinz S, *et al.* Sequence therapy in patients with metastatic renal cell carcinoma: comparison of common targeted treatment options following failure of receptor tyrosine kinase inhibitors. *Eur Urol* 2011;**60**:1163–70	Ineligible intervention
Busch 2013	Busch J, Seidel C, Erber B, Issever AS, Hinz S, Kempkensteffen C, *et al.* Retrospective comparison of triple-sequence therapies in metastatic renal cell carcinoma. *Eur Urol* 2013;**64**:62–70	Ineligible intervention
Calvani 2013	Calvani N, Morelli F, Chiuri V, Gnoni A, Scavelli C, Fedele P, *et al.* Prolonged exposure to tyrosine kinase inhibitors or early use of everolimus in metastatic renal cell carcinoma: are the two options alike? *Med Oncol* 2013;**30**:578	Ineligible intervention
Chen 2012	Chen CC, Hess GP, Liu Z, Gesme DH, Agarwala SS, Garay CC, *et al.* Second-line treatment outcomes after first-line sunitinib therapy in metastatic renal cell carcinoma *Clinical Genitourinary Cancer* 2012;**10**:256–61	Ineligible data
Chen 2012	Chen CC, Hess GP, Liu Z, Gesme DH, Agarwala SS, Hill JW, *et al.* Risk of treatment failure after first-line sunitinib therapy in patients with metastatic renal cell carcinoma. *J Clin Oncol Conf* 2012;**30**(Suppl. 1)	Ineligible data
Clemons 2012	Clemons J, Gao D, Naam M, Breaker K, Garfield D, Flaig TW. Thyroid dysfunction in patients treated with sunitinib or sorafenib. *Clin Genitourin Cancer* 2012;**10**:225–31	Ineligible data

Paper excluded	Full reference details	Reason for exclusion
Derosa 2012	Derosa L, Galli L, Fontana A, Biasco E, Marconcini R, Cianci C, *et al.* Sequential use of treatment options in advanced renal-cell carcinoma (RCC): a retrospective analysis of 42 patient cases. *Ann Oncol* 2012;**23**:ix290	Ineligible data
Di Lorenzo 2011	Di Lorenzo G, Casciano R, Malangone E, Buonerba C, Sherman S, Willet J, *et al.* An adjusted indirect comparison of everolimus and sorafenib therapy in sunitinib-refractory metastatic renal cell carcinoma patients using repeated matched samples. [Erratum appears in *Expert Opin Pharmacother* 2013;**14**:2003], [Erratum appears in *Expert Opin Pharmacother* 2011;**12**:2143]. *Expert Opin Pharmacother* 2011;**12**:1491–7	Ineligible data
Dudek 2009	Dudek AZ, Zolnierek J, Dham A, Lindgren BR, Szczylik C. Sequential therapy with sorafenib and sunitinib in renal cell carcinoma. *Cancer* 2009;**115**:61–7	Ineligible data
Eichelberg 2012	Eichelberg C, Fischer Von Weikersthal L, Goebell P, Lerchenmuller C, Zimmermann U, Freier W, *et al.* Phase III randomised sequential open-label study to evaluate efficacy and safety of sorafenib (SO) followed by sunitinib (SU) vs. sunitinib followed by sorafenib in patients with advanced/meta-static renal cell carcinoma (mRCC) without prior systemic therapy (SWITCH Study) – Safety interim analysis results. *Urologe – Ausgabe A* 2012;**51**:35	Ineligible data
Elaidi 2015	Elaidi R, Harbaoui A, Beuselinck B, Eymard JC, Bamias A, De Guillebon E, *et al.* Outcomes from second-line therapy in long-term responders to first-line tyrosine kinase inhibitor in clear-cell metastatic renal cell carcinoma. *Ann Oncol* 2015;**26**:378–85	Ineligible intervention
Esbah 2014	Esbah O, Demirci U, Helvaci K, Turkoz FP, Ekinci AS, Sonmez OU, *et al.* Sequential therapy and prognostic factors in metastatic renal cell carcinoma: Single centre experience. *J BUON* 2014;**19**:1062–9	Ineligible population
Fischer Von Weikersthal 2012	Fischer Von Weikersthal L, Vervenne WL, Goebell PJ, Eichelberg C, Freier W, De Santis M, *et al.* Phase III randomised sequential open-label study to evaluate efficacy and safety of sorafenib (SO) followed by sunitinib (SU) versus sunitinib followed by sorafenib in patients with advanced/metastatic renal cell carcinoma without prior systemic therapy (SWITCH Study)-Safety interim analysis results. *Onkologie* 2012;**35**:237	Ineligible data
Giuliani 2012	Giuliani J, Drudi F. Immunotherapy and targeted therapies in metastatic renal cell carcinoma: is there a preferred sequence? *Cancer Biother Radiopharm* 2012;**27**:513–18	Ineligible data
Golshayan 2009	Golshayan AR, George S, Heng DY, Elson P, Wood LS, Mekhail TM, *et al.* Metastatic sarcomatoid renal cell carcinoma treated with vascular endothelial growth factor-targeted therapy. *J Clin Oncol* 2009;**27**:235–41	Ineligible data
Grunwald 2011	Grunwald V, Seidel C, Fenner M, Ganser A, Busch J, Weikert S. Treatment of everolimus-resistant metastatic renal cell carcinoma with VEGF-targeted therapies. *Br J Cancer* 2011;**105**:1635–9	Ineligible population
Hahn 2008	Hahn OM, Yang C, Medved M, Karczmar G, Kistner E, Karrison T, *et al.* Dynamic contrast-enhanced magnetic resonance imaging pharmacodynamic biomarker study of sorafenib in metastatic renal carcinoma. *J Clin Oncol* 2008;**26**:4572–8	Ineligible population
Harrison 2013	Harrison MR, George DJ, Walker MS, Chen C, Korytowsky B, Kirkendall DT, *et al.* 'Real world' treatment of metastatic renal cell carcinoma in a joint community-academic cohort: progression-free survival over three lines of therapy. *Clin Genitourin Cancer* 2013;**11**:441–50	Ineligible population
Heng 2011	Heng DYC, Xie W, Bjarnason GA, Vaishampayan U, Tan MH, Knox J, *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	Ineligible data

Paper excluded	Full reference details	Reason for exclusion
Heng 2014	Heng DY, Signorovitch J, Swallow E, Li N, Zhong Y, Wang X, *et al.* Comparative overall survival with treatment sequences for metastatic renal cell carcinoma: a systematic review and meta-analysis of real-world observational studies. *Euro Urol* 2014;**13**:e1142	Ineligible study design
Herrmann 2008	Herrmann E, Bierer S, Gerss J, Kopke T, Hertle L, Wulfing C. Prospective comparison of sorafenib and sunitinib for second-line treatment of cytokine-refractory kidney cancer patients. *Oncology* 2008;**74**:216–22	Ineligible population
Herrmann 2009	Herrmann E, Gerss J, Bierer S, Kopke T, Bolenz C, Hertle L, *et al.* Pre-treatment global quality of health predicts progression free survival in metastatic kidney cancer patients treated with sorafenib or sunitinib. *J Cancer Res Clin Oncol* 2009;**13**:61–7	Ineligible population
Herrmann 2010	Herrmann E, Bierer S, Wulfing C. Update on systemic therapies of metastatic renal cell carcinoma. *World J Urol* 2010;**28**:303–9.	Ineligible study design
Herrmann 2011	Herrmann E, Marschner N, Grimm MO, Ohlmann CH, Hutzschenreuter U, Overkamp F, *et al.* Sequential therapies with sorafenib and sunitinib in advanced or metastatic renal cell carcinoma. *World J Urol* 2011;**29**:361–6	Ineligible data
Hess 2011	Hess GP, Chen C, Liu Z, Gesme DH, Agarwala SS, Hill JW. Risk of treatment failure after first-line tyrosine kinase inhibitors (TKI) therapy in patients with metastatic renal cell carcinoma. *J Clin Oncol* 2011;**29**(Suppl. 1):e15114	Ineligible data
Kontovinis 2012	Kontovinis L, Laschos K, Karadimou A, Andreadis C, Bamias A, Paraskevopoulos P, *et al.* Sequential treatment with sorafenib and sunitinib in metastatic renal cell carcinoma: clinical outcomes from a retrospective clinical study. *Med Oncol* 2012;**29**:750–4	Ineligible study design
La Vine 2010	La Vine DB, Coleman TA, Davis CH, Carbonell CE, Davis WB. Frequent dose interruptions are required for patients receiving oral kinase inhibitor therapy for advanced renal cell carcinoma. *AmJ Clin Oncol* 2010;**33**:217–20	Ineligible population
Latteux 2013	Latteux G, Lebdai S, Hoarau N, Abadie-Lacourtoisie S, Delva R, Chautard D, *et al.* Evaluation of the management of metastatic renal cell carcinoma in the era of targeted therapies. retrospective clinical study over six years. *Prog Urol* 2013;**23**:184–94	Ineligible data
Levy 2013	Levy A, Menard J, Albiges L, Loriot Y, Di Palma M, Fizazi K, *et al.* Second line treatment of metastatic renal cell carcinoma: The Institut Gustave Roussy experience with targeted therapies in 251 consecutive patients. *Eur J Cancer* 2013;**49**:1898–904	Ineligible study design
Li 2014	Li JR, Yang CK, Wang SS, Chen CS, Chiu KY, Cheng CL, *et al.* First-line treatment result influence second-line regimen selection in targeted therapy for metastatic renal cell carcinoma. *Anticancer Res* 2014;**34**:5643–7	Ineligible data
Liu 2009	Liu Z, Zheng J, Riedel AA, Johnson J, Burke J. A retrospective review of treatment discontinuation and survival in patients with advanced renal cell carcinoma treated with sunitinib or sorafenib. *Eur J Cancer* 2009;**7**:429	Ineligible population
Llnassler 2016	Linassler C, Albiges L, Chevreau C, Laguerre B, Oudard S, Gross-Goupll M, *et al.* Everolimus and sunitinib as first- and second-line treatments of patients with metastatic papillary renal cell carcinoma (pRCC): a retrospective study of the GETUG (Groupe Francais d'Etude des Tumeurs Uro-Genitales). *J Clin Oncol* 2016;**34**(Suppl. 2):505	Ineligible data
Macfarlane 2012	Macfarlane R, Heng DY, Xie W, Knox JJ, McDermott DF, Rini BI, *et al.* The impact of kidney function on the outcome of metastatic renal cell carcinoma patients treated with vascular endothelial growth factor-targeted therapy. *Cancer* 2012;**118**:365–70	Ineligible study design
Massard 2010	Massard C, Zonierek J, Gross-Goupil M, Fizazi K, Szczylik C, Escudier B. Incidence of brain metastases in renal cell carcinoma treated with sorafenib. *Ann Oncol* 2010;**21**:1027–31	Ineligible data

Paper excluded	Full reference details	Reason for exclusion
Miyake 2014	Miyake H, Harada K, Inoue TA, Fujisawa M. Assessment of health-related quality of life in Japanese patients with metastatic renal cell carcinoma during treatment with tyrosine kinase inhibitors. *Med Oncol* 2014;**31**:190	Ineligible data
Oudard 2012	Oudard S, Elaidi RT. Sequential therapy with targeted agents in patients with advanced renal cell carcinoma: optimising patient benefit. *Cancer Treat Rev* 2012;**38**:981–7	Ineligible study design
Papavassilis 2014	Papavassilis P, Krabbe LM, Thielen B, Bogemann M, Moritz R, Hoffmeister I, *et al.* Systemic treatment of metastatic renal cell carcinoma: change of paradigms after introduction of targeted therapy. *Der Urologe* 2014;**53**:531–6	Ineligible data
Poprach 2012	Poprach A, Pavlik T, Melichar B, Puzanov I, Dusek L, Bortlicek Z, *et al.* Skin toxicity and efficacy of sunitinib and sorafenib in metastatic renal cell carcinoma: a national registry-based study. *Ann Oncol* 2012;**23**:3137–43	Ineligible population
Poprach 2014	Poprach A, Pavlik T, Melichar B, Kubackova K, Bortlicek Z, Svoboda M, *et al.* Clinical and laboratory prognostic factors in patients with metastatic renal cell carcinoma treated with sunitinib and sorafenib after progression on cytokines. *Urol Oncol* 2014;**32**:488–95	Ineligible population
Procopio 2011	Procopio G, Verzoni E, Iacovelli R, Guadalupi V, Gelsomino F, Buzzoni R. Targeted therapies used sequentially in metastatic renal cell cancer: overall results from a large experience. *Expert Review of Anticancer Therapy* 2011;**11**:1631–40	Ineligible study design
Rini 2008	Rini BI, Choueiri TK, Elson P, Khasawneh MK, Cotta C, Unnithan J, *et al.* Sunitinib-induced macrocytosis in patients with metastatic renal cell carcinoma. *Cancer* 2008;**113**:1309–14	Ineligible population
Risenbeck 2011	Riesenbeck LM, Bierer S, Hoffmeister I, Kopke T, Papavassilis P, Hertle L, *et al.* Hypothyroidism correlates with a better prognosis in metastatic renal cancer patients treated with sorafenib or sunitinib. *World J Urol* 2011;**29**:807–13	Ineligible population
Sablin 2009	Sablin MP, Negrier S, Ravaud A, Oudard S, Balleyguier C, Gautier J, *et al.* Sequential sorafenib and sunitinib for renal cell carcinoma. *J Urol* 2009;**182**:29–34	Ineligible data
Schmidinger 2008	Schmidinger M, Zielinski CC, Vogl UM, Bojic A, Bojic M, Schukro C, *et al.* Cardiac toxicity of sunitinib and sorafenib in patients with metastatic renal cell carcinoma. *J Clin Oncol* 2008;**26**:5204–12	Ineligible population
Schmidinger 2011	Schmidinger M, Vogl UM, Bojic M, Lamm W, Heinzl H, Haitel A, *et al.* Hypothyroidism in patients with renal cell carcinoma. *Cancer* 2011;**117**:534–44	Ineligible data
Schmidinger 2014	Schmidinger M. Improving outcomes in metastatic clear cell renal cell carcinoma by sequencing therapy. *American Society of Clinical Oncology: Educational Book* 2014:e228–38	Ineligible study design
Sherman 2014	Sherman SA, Wang X, Amzal B, Casciano R, Gao H, Stergiopoulos SG, *et al.* A weighted-adjusted indirect comparison of everolimus (EVE) versus axitinib (AXI) in second-line metastatic renal cell carcinoma (mRCC) patients who previously failed sunitinib therapy. *J Clin Oncol* 2014;**32**(Suppl. 4)	Ineligible data
Stenner 2012	Stenner F, Chastonay R, Liewen H, Haile SR, Cathomas R, Rothermundt C, *et al.* A pooled analysis of sequential therapies with sorafenib and sunitinib in metastatic renal cell carcinoma. *Oncology* 2012;**82**:333–40	Ineligible study design
Sun 2013	Sun M, Shariat SF, Trinh QD, Meskawi M, Bianchi M, Hansen J, *et al.* An evidence-based guide to the selection of sequential therapies in metastatic renal cell carcinoma. *Ther Adv Urol* 2013;**5**:121–8	Ineligible study design
Tamaskar 2008	Tamaskar I, Garcia JA, Elson P, Wood L, Mekhail T, Dreicer R, *et al.* Antitumour effects of sunitinib or sorafenib in patients with metastatic renal cell carcinoma who received prior antiangiogenic therapy. *J Urol* 2008;**179**:81–6	Ineligible data

Paper excluded	Full reference details	Reason for exclusion
Vallet 2015	Vallet S, Pahernik S, Hofner T, Tosev G, Hadaschik B, Duensing S, *et al.* Efficacy of targeted treatment beyond third-line therapy in metastatic kidney cancer: Retrospective analysis from a large-volume cancer centre. *Clin Genitourin Cancer* 2015;**13**:e145–e52	Ineligible intervention
Vera-Badillo 2014	Vera-Badillo FE, Templeton A, Ocana A, DeGouveia P, Aneja P, Knox JJ, *et al.* Response to systemic therapy in non-clear cell renal cell carcinomas: A systematic review and meta-analysis. *J Clin Oncol* 2014;**32**(Suppl. 4):425	Ineligible study design
Verzoni 2015	Verzoni E, Grassi P, Montone R, Galli G, Necchi A, Procopio G. TOKIO rationale and protocol: a phase II study to evaluate the activity and safety of third-line tyrosine kinase inhibitor after 2 tyrosine kinase inhibitors in patients with metastatic renal cell carcinoma. *Tumori* 2015;**101**:701–3	Ineligible data
Vickers 2010	Vickers MM, Choueiri TK, Rogers M, Percy A, Finch D, Zama I, *et al.* Clinical outcome in metastatic renal cell carcinoma patients after failure of initial vascular endothelial growth factor-targeted therapy. *Urology* 2010;**76**:430–4	Ineligible intervention
Wagstaff 2016	Wagstaff J, Jones R, Hawkins R, Porfiri E, Pickering L, Bahl A, *et al.* Treatment patterns and clinical outcomes in patients with renal cell carcinoma in the UK: Insights from the RECCORD registry. *Ann Oncol* 2016;**27**:159–65	Ineligible data
Weikert 2011	Weikert S, Seidel C, Busch J, Weinkauf L, Miller K, Gruenwald V. Second-line sequential therapy in patients with metastasized renal cell carcinoma (mRCC): retrospective comparison of common treatment options following failure of receptor tyrosine kinase inhibitor therapy. *J Clin Oncol* 2011;**29**(Suppl. 15):e15059	Ineligible data
Wells 2016	Wells JC, Stukalin I, Norton C, Srinivas S, Lee JL, Donskov F, *et al.* Third-line targeted therapy in metastatic renal cell carcinoma: results from the international metastatic renal cell carcinoma database consortium. *Eur Urol* 2016;**15**:15	Ineligible study design

Economic evaluation

Excluded cost-effectiveness studies

Study	Reference	Reason for exclusion
Anonymous	A randomised controlled trial of IFN-alpha, IL-2 and 5FU versus IFN-alpha alone in metastatic renal cell carcinoma (ISRCTN 46518965). *Urol Oncol* 2002;51–3	Irretrievable
Anonymous	8th Asia Pacific Oncology Summit, APOS 2010. *Japanese J Clin Oncol* 2011;**41**:i1–i18	Irretrievable
Anonymous	Abstracts of the 12th International Kidney Cancer Symposium. *BJU Int* 2013;**112**(Suppl. 3):1–17	Abstract with insufficient methodological details
Alam 2012	Alam M, Delahoy P, Park SH. Progression free survival vs overall survival: an example from randomised phase III trial with axitinib (AXIS) in metastatic renal cell carcinoma. Asia-Pacific Journal of Clinical Oncology Conference; 2012;**8**:104–14	Abstract with insufficient methodological details
Alexander 2007	Alexander W. Renal cancer. *Pharmacy and Therapeutics* 2007;**32**:680	Irretrievable
Anaya 2012	Anaya P, Delea TE, Pichardo P, Diaz JR. Cost-effectiveness analysis based on progression free survival (PFS) of pazopanib versus sunitinib for the treatment of advanced renal cell carcinoma (ARCC) in the Mexican context. *Value Health* 2012;A220	Abstract with insufficient methodological details

Study	Reference	Reason for exclusion
Arreola-Ornelas 2011	Arreola-Ornelas H, Rosado-Buzzo A, Garcia-Mollinedo M, Camacho L, Mould-Quevedo JF, Muciño-Ortega E, Galindo-Suarez RM. Cost-effectiveness of temsirolimus for metastic renal-cell carcinoma and poor prognosis patients in Mexico. *Value Health* 2011;A164	Abstract with insufficient methodological details
Atzpodien 1996	Atzpodien J. Interleukin 2 based ambulatory therapy of metastatic renal cell carcinoma. *Medizinische Klinik* 1996;(Suppl. 3):38–43	Irretrievable
Ballali 2013	Ballali S, Chiffi D, Trojniak MP, Gregori D. Economic impact of sunitinib and sorafenib use in metastatic renal cell carcinoma treatment in Veneto Region, Italy. *Open Pharm J* 2013;**7**:2–8	Non-UK costing study
Barber 2010	Barber J, M Button, C Jones, K Das, B Amphlett, S Kumar, J Lester, J Tanguay. mTOR inhibition and renal cell carcinoma; a comparison between sirolimus and temsirolimus. *Ann Oncol* 2010;**21**(Suppl. 8):viii 288	Abstract with insufficient methodological details
Barbosa 2013	Barbosa MMA, Almeida AM, Costa JDO, Júnior AG, Acurcio FA. Sorafenib for kidney cancer: evidence of efficacy, safety and cost estimates. *Value Health* 2013;A128–9	Abstract with insufficient methodological details
Benedict 2011	Benedict A, Figlin R, Sandström P, Harmenberg U, Ullén A, Charbonneau C, *et al.* Economic evaluation of new targeted therapies for the first-line treatment of patients with metastatic renal cell carcinoma. *BJU International* 2011;**108**:665–72	Wrong population
Benedict 2015	Benedict A, Ramaswamy K, Sandin R. Cost-effectiveness of pazopanib versus sunitinib for renal cancer in the United States. *Journal of Managed Care & Specialty Pharmacy* 2015;**21**:834–40	Letter/commentary
Benedict 2008	Benedict A, Charbonneau C, Kim ST, Negrier S. Cost-effectiveness of sunitinib (SU), sorafenib (SFN), temsirolimus (TMS), and bevacizumab plus interferon-alfa (BEV/IFN) as 1st-line therapy for metastatic renal cell carcinoma (MRCC) – an indirect comparison. *Ann Oncol* 2008;**19**(Suppl. 8):viii 227	Abstract with insufficient methodological details
Benedict 2009	Benedict A, Figlin R, Charbonneau C, Kreif N, Hariharan S, Négrier S. Economic evaluation of sunitinib versus other new targeted therapies as first-line treatment of metastatic renal cell carcinoma (mRCC) in the United States. *J Clin Oncol* 2009;e17556	Abstract with insufficient methodological details
Berghea 2013	Berghea F, Skoupa J, Ciuleanu T, Miron L, Stanculeanu DL, Jinga D, *et al.* A cost-effectiveness analyses of using sunitinib (SU) in first line of metastatic renal cancer in Romanian jurisdiction. *Value Health* 2013;A411–12	Abstract with insufficient methodological details
Bodnar 2012	Bodnar C, Paramore LC, Knopf KB. Economic evaluation of reduced futile 1st line therapy in metastatic renal cell carcinoma patients using early angiogenesis-specific imaging. *Value Health* 2012;A353	Abstract with insufficient methodological details
Bonastre 2009	Bonastre J, Chevalier J, Koscielny S, Lassau N. Dynamic contrast-enhanced ultrasound with quantification to assess targeted treatment efficacy: results of a multi-centric prospective cost study. *Value Health* 2009;A281	Abstract with insufficient methodological details
Bonthapally 2009	Bonthapally V, Ghosh. Cost-effectiveness and budget impact analysis of using temsirolimus compared to interferon alpha in metastatic renal cell carcinoma. *Value Health* 2009;A263	Abstract with insufficient methodological details
Borker 2014	Borker R. Costs associated with adverse events in patients with metastatic renal cell carcinoma. *J Med Econ* 2014;**17**:792–7	Non-UK costing study
Borovicka 2010	Borovicka JH, Mulcahy M, Calahan C, Lacouture ME. Economic impact in the management of dermatological toxicities (DTS) induced by multikinase inhibitors: sorafenib and sunitinib in renal cell carcinoma (RCC). *Supportive Care in Cancer* 2010;**18**(Suppl.3):67	Abstract with insufficient methodological details

Study	Reference	Reason for exclusion
Calderero 2009	Calderero, García-Muro X, Puente J, Trigo JM, Castro AJ, Martín-Escudero V, Yébenes M, *et al.* Cost of managing adverse events in the treatment of first line metastasic renal cell carcinoma: bevacizumab + interferon alpha-2 A compared with sunitinib in Spain. *Value Health* 2009;A263	Wrong population
Calvo Aller 2011	Calvo Aller E, Maroto P, Kreif N, Larriba JLG, López-Brea M, Castellano D, *et al.* Cost-effectiveness evaluation of sunitinib as first-line targeted therapy for metastatic renal cell carcinoma in Spain. *Clin Trans Oncol* 2011;**13**:869–77	Wrong population
Cardona 2009	Cardona, AF, Caceres HA, Spath A, Lujan M, Lopera D, Otero JM, Carranza H, *et al.* Budgetary impact of metastatic renal cell carcinoma (MRCC) treatment on the Colombian general health social security system (SGSSS). *Value Health* 2009;A44	Abstract with insufficient methodological details
Carlos 2010	Carlos F, Ramirez J, Aguirre A. Costs of managing adverse events of first-line therapy for metastatic renal cell carcinoma in mexico: bevacizumab in combination with interferon-alpha-2a compared with sunitinib. *Value Health* 2010;A258–9	Abstract with insufficient methodological details
Casciano 2011	Casciano R, Chulikavit M, Di Lorenzo G, Liu Z, Baladi JF, Wang X, *et al.* Economic evaluation of everolimus versus sorafenib for the treatment of metastatic renal cell carcinoma after failure of first-line sunitinib. *Value Health* 2011;**14**:846–51	Duplicate
Casciano 2010	Casciano R, Chulikavit, Di Lorenzo G, Stern L, Liu Z, Wang X, Garay C, Garrison L. Economic evaluation of everolimus versus sorafenib for the treatment of advanced renal cell carcinoma after failure on treatment with sunitinib. *Ann Oncol* 2010;viii295	Abstract with insufficient methodological details
Casciano 2010	Casciano R, Chulikavit M, Zheng J, Liu Z, Rogerio J. Estimated impact of everolimus on annual drug expenditure in the treatment of advanced renal cell carcinoma in a US health plan. *Value Health* 2010;A28	Abstract with insufficient methodological details
Casciano 2011	Casciano R, Chulikavit M, El Ouagari K, Wang X. Cost-effectiveness of treating metastatic renal cell carcinoma (MRCC) patients whose disease failed on one prior VEGF-TKI therapy with everolimus compared to treating with best supportive care (BSC) alone in Canada. *Value Health* 2011;A445–6	Abstract with insufficient methodological details
Castellano 2009	Castellano D, De la Rosa F, Rodriguez Antolin A, Villacampa F, Sepulveda J, Ghanem I, *et al.* Sunitinib therapy for patients with Advanced Renal Cell Carcinoma (ARCC): analysis for safety and activity on single institution experience: favourable overall survival according MSKCC-group risk. *Urology* 2009:**74**:S113	Abstract with insufficient methodological details
Castellano 2008	Castellano R, Sepulveda J, Coronado C, Garcia Rodriguez L, Garcia Escobar I, Diaz Padilla I, Sepulveda D. Sunitinib therapy for patients with advanced renal cell carcinoma (ARCC): analysis for safety and activity on single institution experience. Favourable overall survival according MSKCC group risk. *Ann Oncol* 2008;**19**:viii193	Abstract with insufficient methodological details
Chabot 2010	Chabot I, Rocchi A. How do cost-effectiveness analyses inform reimbursement decisions for oncology medicines in Canada? The example of sunitinib for first-line treatment of metastatic renal cell carcinoma. *Value Health* 2010;**13**:837–45	Wrong population
Chandiwana 2014	Chandiwana D, Perrin A, Sherman S. A cost effectivness analysis of everolimus compared with axitinib in the treatment of metastatic renal cell carcinoma in the United Kingdom. *Value Health* 2014;A640	Abstract with insufficient methodological details
Charalambous 2010	Charalambous H, Chulikavit M. Metastatic renal cell carcinoma (mRCC): the budget impact for the introduction of everolimus (E) in the Cyprus health care system. *Ann Oncol* 2010;**21**	Abstract with insufficient methodological details
Chen 2010	Chen CG. Observational study evaluating resource utilisation among metastatic renal cell carcinoma patients treated with mTOR inhibitors in the outpatient community-based setting. *J Clin Oncol* 2010;**28**(Suppl.15):e15027	Abstract with insufficient methodological details

Study	Reference	Reason for exclusion
Chen 2010	Chen KS. Health care costs associated with angiogenesis inhibitors (AIS) and mtor inhibitors (MTORS) in patients with metastatic renal cell carcinoma (MRCC) treated at us community oncology clinics. *Value Health* 2010;A33	Abstract with insufficient methodological details
Choueiri 2012	Choueiri TK. Costs associated with angiogenesis inhibitor therapies for metastatic renal cell carcinoma in clinical practice: results from a medical chart review study. *Urol Oncol* 2012;**30**:848–55	Non-UK costing study
De Groot 2013	De Groot S, Blommestein H, Redekop W, Oosterwijk E, Kiemeney L, Uyl- de Groot C. The cost-effectiveness of sequential first- and second-line treatments in metastatic renal cell carcinoma using real-world data and a patient-level simulation model. *Value Health* 2013;**16**:A587	Abstract with insufficient methodological details
De Groot 2012	DeGroot S, Redekop W. The evaluation of the use and effectiveness of bevacizumab for patients with metastatic renal cell carcinoma in daily practice. *Value Health* 2012:A409–10	Abstract with insufficient methodological details
Delea 2015	Delea TE, Amdahl J, Diaz J, Nakhaipour HR, Hackshaw MD. Cost-effectiveness of pazopanib versus sunitinib for renal cancer in the United States. *J Manag Care Pharm* 2015;**21**:46–54	Wrong population
Delea 201	Delea TE, Amdahl J, Diaz J, Nakhaipour HR, Hackshaw MD. Cost-effectiveness of pazopanib versus sunitinib for renal cancer in the United States. *J Manag Care Pharm* 2015;**21**:46–54	Duplicate
Demlova 2009	Demlova R, Ondrackova B, Kominek J. The economic evaluation of sunitinib and sorafenib in MRCC patients in the Czech Republic. *Value Health* 2009;A266	Abstract with insufficient methodological details
Dial 2009	Dial E, Duh M, Fournier A, Antras L, Rodermund D, Neary MP, Oh WK. Cost implications of intravenous bevacizumab treatment in patients with renal cell carcinoma (RCC): a retrospective claims database analysis. *J Clin Oncol* 2009;**27**:5112	Abstract with insufficient methodological details
Dial 2010	Dial E, Duh MS, Antras L, Rodermund D, Neary MP, Choueiri TK, Oh WK. Incidence and cost of adverse events (AES) in patients with renal cell carcinoma (RCC) treated with angiogenesis inhibitors (AIS). *Value Health* 2010; A76	Abstract with insufficient methodological details
Diaz 2008	Diaz S, Calvo Aller E, Maroto P, Puente J, Lopez-Brea M, Castellano D. Cost-effectiveness and cost–utility analysis of sunitinib (SU) vs sorafenib (SFN) and bevacizumab 1 interferon-alfa (BEV/IFN) as first-line treatment for metastatic renal cell carcinoma (MRCC) in Spain. *Ann Oncol* 2008;**19**(Suppl. 8):viii227	Abstract with insufficient methodological details
Duh 2009	Duh MS, Dial E, Choueiri TK, Fournier AA, Antras L, Rodermund D, *et al.* Cost implications of IV versus oral anti-angiogenesis therapies in patients with advanced renal cell carcinoma: retrospective claims database analysis. *Curr Med Res Opin* 2009;**25**:2081–90	Non-UK costing study
Ebara 2013	Ebara T, Ohno T, Nakano T. Quantitative medical cost-effectiveness analysis of molecular-targeting cancer drugs in Japan. *Daru* 2013;**21**:40	Wrong population
El-Ougari 2010	El Ouagari, K. Chulikavit. Cost-Effectiveness of Treating Metastatic Renal Cell Carcinoma (mRCC) Patients Whose Disease Failed on VEGF-TKI Therapies with Everolimus Compared to Treating with Best Supportive Care (BSC) Alone: a Canadian Societal Perspective. 35th ESMO Congress, Milan, 8–12 October 2010	Abstract with insufficient methodological details
Elsisi 2014	Elsisi GH. Cost-effectiveness of pazopanib versus sunitinib in Egyptian patients with metastatic renal cell carcinoma from the health insurance perspective: A Markov model. *Value Health* 2014;A90–1	Abstract with insufficient methodological details
Espinosa 2014	Espinosa JG-L. Cost-utility analysis of pazopanib versus sunitinib as first-line treatment of metastatic renal cell carcinoma (mRCC) in Spain. *Value Health* 2014;A632–3	Abstract with insufficient methodological details

Study	Reference	Reason for exclusion
Ferreira 2014	Ferreira CNR. Cost analysis of adverse events associated with first line treatment for metastatic renal cell carcinoma (MRCC) in the perspective of public and private health insurance in Brazil. *Value Health* 2014;A76	Abstract with insufficient methodological details
Ferreira 2015	Ferreira CNS. Economic burden of adverse events associated with first line metastatic renal cell carcinoma (MRCC) treatment in public and private Brazilian perspective. *Value Health* 2015;A251	Abstract with insufficient methodological details
Geynisman 2013	Geynisman DMH. Adherence (ADH) to and beliefs about oral anticancer medications (OAMs) in patients (PTS) with metastatic renal cell carcinoma (mRCC). *J Clin Oncol* 2013;**31**:25	Abstract with insufficient methodological details
Geynisman 2015	Geynisman DMH. Treatment patterns and costs for metastatic renal cell carcinoma patients with private insurance in the United States. *Clin Genitourin Cancer* 2015;**13**:e93–e100	Non-UK costing study
Geynisman 2014	Geynisman DMS. Treatment (tx) patterns and drug (Rx) costs for patients (pts) with metastatic renal cell carcinoma (mRCC) in the United States. *J Clin Oncol* 2014;**32**(Suppl.4):457	Abstract with insufficient methodological details
Godoy 2009	Godoy JC. Cost-effectiveness analysis of first-line treatment for metastatic renal cell carcinoma (mRCC) in Colombia (ONCOLGroup study). *J Clin Oncol* 2009;e16150	Abstract with insufficient methodological details
Godoy 2009	Godoy JIC. Cost-effectiveness analysis of first-line treatment for metastatic renal cell carcinoma (MRCC) in Colombia (ONCOLGROUP STUDY). *Value Health* 2009;A495	Abstract with insufficient methodological details
Greenberg 2009	Greenberg D. Economic evaluation of Sunitinib, Sorafenib, Bevacizumab/interferon alpha and Temsirolimus in first line treatment of metastatic renal cell carcinoma in Israel. *Value Health* 2009;A42	Abstract with insufficient methodological details
Gruschkus 2012	Gruschkus SKB. Avoidance of futile treatment and adverse events by using angiogenesis-specific imaging for early detection of disease progression in patients with first-line metastatic renal cell carcinoma. *J Clin Oncol* 2012;**30**(Suppl.34):134	Abstract with insufficient methodological details
Hackshaw 2014	Hackshaw MDH. Costs associated with health care resource use in patients with advanced renal cell carcinoma receiving first-line treatment with pazopanib versus sunitinib. *Value Health* 2014;A77–8	Abstract with insufficient methodological details
Hagiwara 2013	Hagiwara M. Economic burden of selected adverse events in patients aged > 65 years with metastatic renal cell carcinoma. *J Med Econ* 2013;**16**:1300–6	Wrong population
Hagiwara 2013	Hagiwara M, Hagiwara M, Borker R, Oster G. Economic burden of adverse events in patients with metastatic renal cell carcinoma. *Clin Ther* 2013;**35**:1955–63	Non-UK costing study
Hanninen 1996	Hanninen EL. Interleukin-2 based home therapy of metastatic renal cell carcinoma: risks and benefits in 215 consecutive single institution patients. *J Urol* 1996;**155**:19–25	Wrong study type
Hansen 2015	Hansen RN. Health care costs among renal cancer patients using pazopanib and sunitinib. *J Manag Care Spec Pharm* 2015;**21**:37–44	Non-UK costing study
Harnett 2015	Harnett JM. Sunitinib and pazopanib treatment patterns and cost outcomes in Medicare supplemental-covered patients with renal cell carcinoma. *J Clin Oncol* 2015;**33**(Suppl.7):485	Abstract with insufficient methodological details
Harnett 2015	Harnett JM. Treatment patterns and costs associated with sunitinib and pazopanib treatment for renal cell carcinoma: a commercial health claims analysis. *Value Health* 2015;A196–7	Abstract with insufficient methodological details
Henk 2013	Henk HJ. Retrospective claims analysis of best supportive care costs and survival in a US metastatic renal cell population. *Clinicoecon Outcome Res* 2013;**5**:347–54	Non-UK costing study
Henriksson 1998	Henriksson R. Survival in renal cell carcinoma-a randomised evaluation of tamoxifen vs interleukin 2, alpha-interferon (leucocyte) and tamoxifen. *Br J Cancer* 1998;**77**:1311–7	No relevant outcomes

Study	Reference	Reason for exclusion
Hoyle 2010	Hoyle M, Green C, Thompson Coon J, Liu Z, Welch K, *et al.* Cost-effectiveness of temsirolimus for first line treatment of advanced renal cell carcinoma. *Value Health* 2010;**13**:61–8	Wrong population
James 2009	James NP. Effect of the UK postcode lottery on survival of patients with metastatic renal cancer: an audit of outcomes in patients with metastatic renal cancer suitable for treatment with tyrosine kinase inhibitors. *Clin Oncol* 2009;**21**:610–16	Wrong study type
Jirillo 2012	Jirillo A. The impact of new cancer drugs in real practice oncology: A monoinstitutional experience. *Immunopharmacol Immunotoxicol* 2012;**34**:702–5	Wrong study type
Jones 2011	Jones CB. mTOR inhibition and renal cell carcinoma: A comparison between sirolimus and temsirolimus. *Clin Oncol* 2011;70	Irretrievable
Kan 2012	Kan HCP. Cost-effectiveness analysis of temsirolimus in patients with poor risk renal-cell carcinoma. *Value Health* 2012;20120902	Abstract with insufficient methodological details
Kim 2014	Kim SP. Out-of-pockets costs for patients receiving targeted agents for metastatic renal cell carcinoma. *J Clin Oncol* 2014;**32**(Suppl.15):e15598	Abstract with insufficient methodological details
Kolbin 2015	Kolbin A. Pharmacoeconomic analysis of the use of everolimus compared to axitinib in second line therapy of patients with metastatic renal cell carcinoma. *Value Health* 2015;**18**:A442	Abstract with insufficient methodological details
Kostyuk 2014	Kostyuk A. Cost-effectiveness evaluation of sunitinib as first-line targeted therapy for metastatic renal cell carcinoma in Kazakhstan. *Value Health* 2014;A85	Abstract with insufficient methodological details
Kovacs 2012	Kovacs E, Nagy B, Bidlo J. Medical management of metastatic renal cell carcinoma, retrospective analysis of real world data settings. *Value Health* 2012;A412	Abstract with insufficient methodological details
Kulikov 2014	Kulikov A, Kumirov I. Pharmacoeconomic analysis of axitinib as second-line treatment for metastatic renal cell carcinoma. *Value Health* 2014;A638–9	Abstract with insufficient methodological details
Liu 2010	Liu Z. Comparative outpatient resource utilisation study of metastatic renal cell carcinoma patients receiving oral vs. intravenous mtor inhibitors. *Ann Oncol* 2010;**21**(Suppl.8):VIII286	Abstract with insufficient methodological details
Luo 2009	Luo X. Using the rasch model to validate and enhance the interpretation of the functional assessment of cancer therapy-kidney symptom index – disease-related symptoms scale. *Value Health* 2009;**12**:580–6	No relevant outcomes
Margolis 2015	Margolis J, Princic N, Doan J, Lenhart G, Motzer R. Cost comparison of first line metastatic renal cell carcinoma treatments using a retrospective claims dataset. *Value Health* 2015;A197	Abstract with insufficient methodological details
Martín-Vila 2012	Martín-Vila A, Ivarez Seoane JA, Pérez Parente D, Ivarez-Payero MA, Ucha Samartin M, Martínez-López de Castro N. Sunitinib in advanced/metastatic renal cell carcinoma in adults. *Int J Clin Pharm* 2012;**34**:239	Abstract with insufficient methodological details
Martín-Vila 2013	Martín-Vila A, Ivarez-Payero MA, Martínez-López de Castro N, Suáez-Santamaría M, Castro-Domínguez JM, Ascunce-Saldaña MDP. Everolimus in advanced/metastatic renal cell carcinoma in adults after sunitinib progression. *Int J Clin Pharm* 2013;**35**:907	Abstract with insufficient methodological details
Matusewicz 2010	Matusewicz W, Baran J, Farkowski MM. Utilising evidence from different levels in the reimbursement process of new medical technologies-advanced renal cell carcinoma first line therapy in Poland 2008–2009. *Value Health* 2010;A217	Abstract with insufficient methodological details
Mazelova 2014	Mazelova J. Reimbursed pharmacotherapy of metastatic clear cell kidney cancer (mCCKC) in the Czech Republic. *Value Health* 2014;A658	Abstract with insufficient methodological details
Mei 2009	Mei S. Cost implications of IV versus oral anti-angiogenesis therapies in patients with advanced renal cell carcinoma: Retrospective claims database analysis. *Curr Med Res Opin* 2009;**25**:2081–90	Duplicate

Study	Reference	Reason for exclusion
Mickisch 2010	Mickisch G. Costs of managing adverse events in the treatment of first-line metastatic renal cell carcinoma: bevacizumab in combination with interferon-alpha2a compared with sunitinib. *Br J Cancer* 2010;**102**:80–6	Non-UK costing study
Mickisch 2010	Mickisch G. Costs of managing adverse events in the treatment of first-line metastatic renal cell carcinoma: bevacizumab in combination with interferon-alpha2a compared with sunitinib. *J Urol* 2010;**184**:1303–4	Duplicate
Mickisch 2009	Mickisch G. Cost of managing side effects of first-line therapy for metastatic renal cell carcinoma (MRCC) in Germany, France, UK and Italy: bevacizumab (BEV) interferon-alpha2a compared with sunitinib. *Value Health* 2009;A39	Abstract with insufficient methodological details
Mickisch 2009	Mickisch G. Cost of managing side effects of first-line bevacizumab (BEV) lower-dose interferon-alpha2a in patients with metastatic renal cell carcinoma (MRCC) in Germany, France, and United Kingdom. *Value Health* 2009;A38–9	Duplicate
Miguel 2014	Miguel L. Economic evaluation of axitinib for second line treatment in adult patients with advanced renal cell carcinoma-the Portuguese case. *Value Health* 2014;A639–A4	Abstract with insufficient methodological details
Mihajlovic 2013	Mihajlovic J, Minovic I. Cost utility analysis of everolimus in the treatment of metastatic renal cell cancer in the Netherlands. *Value Health* 2013;A416	Abstract with insufficient methodological details
Mihajlovic 2012	Mihajlovic J. Cost-effectiveness of everolimus for second line treatment of metastatic renal cell cancer in Serbia. *Value Health* 2012;A427	Abstract with insufficient methodological details
Mihajlovic 2014	Mihajlovic J. Cost-effectiveness of targeted therapeutics in metastatic renal cell cancer seen from two different economic perspectives. *Value Health* 2014;A90	Abstract with insufficient methodological details
Moyneur 2010	Moyneur E, Dorff TB, Barghout V, Meyers S, Hu J, Quinn DI. Retrospective claims database cost analysis of second-line sorafenib (SR) or sunitinib (SR) therapy in treatment of patients (pts) with renal cell carcinoma (RCC). *J Clin Oncol* 2010;**28**(Suppl.15):e16521	Abstract with insufficient methodological details
Moyneur 2010	Moyneur E. Retrospective US claims database analysis of the cost of sequencing of sorafenib and sunitinib in the treatment of patients with renal cell carcinoma (RCC). *Value Health* 2010;**13**. Conference A33.	Abstract with insufficient methodological details
Munir 2008	Munir U. Cost-effectiveness of sunitinib (SU) vs sorafenib (SFN), temsirolimus (TMS) and bevacizumab 1 interferon-alfa (BEV/IFN) as first-line therapy for metastatic renal cell carcinoma (MRCC) – adaptation for the Swedish health service. *Ann Oncol* 2008;**19**(Suppl.8):viii225–8	Abstract with insufficient methodological details
Nakhaipour 2014	Nakhaipour HR. Cost-effectiveness of pazopanib (PAZ) versus sunitinib (SUN) as first-line treatment of metastatic renal cell carcinoma (mRCC) patients in the United States. *J Clin Oncol* 2014;**32**	Abstract with insufficient methodological details
Negrier 2000	Negrier S. Cytokine treatment for metastatic renal carcinoma – the experience of the Groupe Francais d'Immunotherapie. *J Clin Oncol* 2000:28–33	Irretrievable
Negrier 2012	Negrier S. Interpreting overall survival (OS) results when progression free survival (PFS) benefits exists in today's oncology landscape-metastatic renal cell carcinoma (MRCC) case study. *Ann Oncol* 2012;**23**(Suppl.9):ix453	Abstract with insufficient methodological details
Nevarez-Sida 2012	Nevarez-Sida A. Economic evaluation of everolimus as second line treatment in metastatic renal cancer in Mexico. *Value Health* 2012; A221	Abstract with insufficient methodological details
Nunez 2012	Nunez S. Cost-effectiveness analysis of first-line treatment for metastatic renal cell carcinoma in Colombia. *Value Health* 2012;A219	Abstract with insufficient methodological details

Study	Reference	Reason for exclusion
Oh 2010	Oh W. Costs of treatment with angiogenesis inhibitors (AIS) in patients with metastatic renal cell carcinoma (MRCC): Results from a medical chart review study. *Value Health* 2010;A32	Abstract with insufficient methodological details
Ondrackova 2010	Ondrackova B. Economic evaluation of targeted biologic therapy in metastatic renal cell carcinoma. *Klinicka Onkologie* 2010;**23**:439–45	Study not available in English language
Ondrackova 2010	Ondrackova BD. Sorafenib and sunitinib in metastatic renal cell carcinoma: cost-effectiveness analysis in reimbursement proceedings vs. data from clinical practice. *Value Health* 2010;A267	Abstract with insufficient methodological details
Ozer-Stillman 2013	Ozer-Stillman I. An economic analysis of axitinib and sorafenib for second-line treatment of cytokine-refractory patients with advanced renal cell carcinoma in the United States (US). *J Clin Oncol* 2013;**31**(Suppl.15):e15601	Abstract with insufficient methodological details
Ozer-Stillman 2012	Ozer-Stillman I, Keyser R, Ambavane A, Cislo P. Sorafenib versus axitinib for second-line treatment of sunitinib-refractory patients with advanced renal cell carcinoma (RCC) in the United States (US). *Ann Oncol* 2012;**23**(Suppl.9):ix258–93	Abstract with insufficient methodological details
Paglino 2010	Paglino C. Health care costs associated with multikinase inhibitors (MKIS) for treatment of metastatic renal cell carcinoma (MRCC) in a clinical practice setting in Italy. *Value Health* 2010;A33	Abstract with insufficient methodological details
Perrin 2013	Pal S, Perrin S. The lifetime cost of everolimus vs. axitinib in metastatic renal cell carcinoma patients who have failed prior sunitinib therapy in the US. *BJU Int* 2013;**112**(Suppl.3):3	Abstract with insufficient methodological details
Park 2012	Park MH, Jo C, Bae EY, Lee EK. A comparison of preferences of targeted therapy for metastatic renal cell carcinoma between the patient group and health care professional group in South Korea. *Value Health* 2012;**15**:933–39	No relevant outcomes
Pepe 2012	Pepe C, Paladini L, Sedlmayer C, Machado M. Cost-effectiveness of pazopanib versus sunitinib and bevacizumab associated to interferon alpha as first line treatments for metastatic renal cell carcinoma. *Value Health* 2012;A218	Abstract with insufficient methodological details
Pepe 2013	Pepe C, Sedlmayer C, Machado M. Cost-effectiveness of pazopanib as first line treatment for metastatic renal cell carcinoma in Brazil: updated analysis. *Value Health* 2013;A685	Abstract with insufficient methodological details
Perrin 2015	Perrin A, Sherman S, Pal S, Chua A, Gorritz M, Liu Z, *et al.* Lifetime cost of everolimus vs axitinib in patients with advanced renal cell carcinoma who failed prior sunitinib therapy in the US. *J Med Econ* 2015;**18**:200–9	Non-UK costing study
Perrin 2013	Perrin A, Chua A, Wang X, Hurry M. Cost of care with everolimus versus axitinib for second-line metastatic renal cell carcinoma patients in Canada. *Value Health* 2013;A402	Abstract with insufficient methodological details
Piga 1997	Piga A. A phase II study of interferon alpha and low-dose subcutaneous interleukin-2 in advanced renal cell carcinoma. *Cancer Immunol Immunother* 1997;**44**:348–51	No relevant outcomes
Procopio 2008	Procopio G, Verzoni E, Bajetta E, Giuliani G, Peccerillo C, Walzer S, Nuijten M. Costs of managing side effects in the treatment of first line metastatic renal cell carcinoma (MRCC) in Italy: bevacizumab (BEV) + interferon ALPHA-2 A (IFN) compared with sunitinib. *Ann Oncol* 2008;**19**(Suppl.8):viii187–207	Abstract with insufficient methodological details
Puento 2009	Puente J, Calderero V. Costs of adverse events management associated to the treatment of first-line metastatic renal cell carcinoma with bevacizumab + interferon alpha-2a compared with sunitinib in Spain. *Eur J Cancer* 2009;436	Abstract with insufficient methodological details
Purmonen 2008	Purmonen T, Nuttunen P, Vuorinen R, Pyrhönen S, Kataja V, Kellokumpu-Lehtinen P. Cost and survival analysis of interferon treatment in metastatic renal cell carcinoma. *Ann Oncol* 2008;**19** (Suppl.8):viii187–207	Abstract with insufficient methodological details

Study	Reference	Reason for exclusion
Purmonen 2009	Purmonen T, Vuorinen R, Kataja V, Pyrhönen S, Kellokumpu-Lehtinen P. Cost of renal cell carcinoma treatment in patients treated with interferon-alpha. *Value Health* 2009;A266–7	Abstract with insufficient methodological details
Purmonen 2010	Purmonen T, Nuttunen P, Vuorinen R, Pyrhönen S, Kataja V, Kellokumpu-Lehtinen P. Current and predicted cost of metastatic renal cell carcinoma in Finland. *Acta Oncologica* 2010;**49**:837–43	Non-UK costing study
Quinn 2009	Quinn D, Barghout V, Moyneur E. Medical costs of sorafenib compared with sunitinib in treatment of patients < 65 years with renal cell carcinoma: A retrospective claims database analysis. *J Clin Oncol* 2009:e17536	Abstract with insufficient methodological details
Quinn 2009	Quinn, D. J. B. Retrospective claims database analysis of the direct medical costs associated with sorafenib and sunitinib in the treatment of patients with renal cell carcinoma who are under 65 years old. *Value Health* 2009:A41	Abstract with insufficient methodological details
Racsa 2015	Racsa PN, Whisman TR, Worley K. Comparing two tyrosine kinase inhibitors for treatment of advanced renal cell carcinoma in Medicare and commercially insured patients. *Curr Med Res Opin* 2015;**31**:1933–40	Non-UK costing study
Ramirex 2012	Ramírez MA, Peniche G, Rodríguez JA, Nuño-Langre C, Muciño-Ortega E, Mould-Quevedo JF. Economic analysis of adverse events produced by sunitinib, interferon alpha and bevacizumab + interferon alpha in patients with metastatic renal cell cancer in Mexico. *PharmacoEconomics* 2012;**9**:145–57	Study not available in English language
Ravasio 2011	Ravasio R, Ortega C, Sabbatini R, Porta C. Bevacizumab plus interferon-alpha versus sunitinib for first-line treatment of renal cell carcinoma in Italy: a cost-minimisation analysis. *Clin Drug Invest* 2011;**31**:507–17	Wrong population
Ravasio 2012	Ravasio R, Ortega C, Sabbatini R, Porta C. Bevacizumab plus interferon-a versus sunitinib for first-line treatment of renal cell carcinoma in Italy. a cost-minimisation analysis. *PharmacoEconomics* 2012;**14**:150–1	Study not available in English language
Remak 2008	Remak E, Charbonneau C, Negrier S, Kim ST, Motzer RJ. Economic evaluation of sunitinib malate for the first-line treatment of metastatic renal cell carcinoma. *J Clin Oncol* 2008;**26**:3995–4000	Wrong population
Remak 2009	Remak E, Vioix H, Sandin R, Harmenberg U, Ullen A, Sandstrom P. Cost-effectiveness analysis of sunitinib, bevacizumab + interferon-alfa and temsirolimus as first-line therapy of metastatic renal cell carcinoma in Sweden. *Value Health* 2009:A270	Abstract with insufficient methodological details
Rosado-Buzzo 2015	Rosado-Buzzo A, Albuja M, Garcia-Molliendo L, Luna-Casas G. Economic impact of the addition of axitinib as a second line treatment for metastatic renal cell carcinoma in the Ecuatorian public heathcare sector. *Value Health* 2015:A195–6	Abstract with insufficient methodological details
Salinas-Escudero 2009	Salinas-Escudero G, Contreras-Hernandez I, Mould-Quevedo J. Cost-effectiveness and cost–utility analysis of sunitinib vs sorafenib and bevacizumab + interferon-alfa as firstline treatment for metastatic renal cell carcinoma in Mexico. *Value Health* 2009:A497	Abstract with insufficient methodological details
Shi 2014	Shi Q, Yin H, Xuan J, Wu Y, Cheng G. Cost-effectiveness of sunitinib as first-line targeted therapy for metastatic renal cell carcinoma in China. *Value Health* 2014:A638	Abstract with insufficient methodological details
Shih 2011	Shih YC, Chien CR, Xu Y, Pan IW, Smith GL, Buchholz TA. Economic burden of renal cell carcinoma in the US: part II-an updated analysis. *PharmacoEconomics* 2011;**29**:331–41	Non-UK costing study
Silva 2010	Silva C, Monteiro I, Schwander B. Cost analysis of managing adverse events in the treatment of metastatic renal cell carcinoma in Portugal: a comparison of bevacizumab in combination with interferon alfa-2a and sunitinib. *Value Health* 2010:A258	Abstract with insufficient methodological details

Study	Reference	Reason for exclusion
Silverio 2009	Silverio N, Yang S, Alemao E. Cost-effectiveness analysis of temsirolimus vs. sunitinib malate in poor prognosis metastatic renal cell carcinoma (MRCC) in Portugal. *Value Health* 2009:A271	Abstract with insufficient methodological details
Soerensen 2015	Soerensen AV, Donskov F, Kjellberg J, Ibsen R, Hermann GG, Jensen NV, *et al.* Health economic changes as a result of implementation of targeted therapy for metastatic renal cell carcinoma: national results from DARENCA study 2. *Eur Urol* 2015;**68**:516–22	Non-UK costing study
Sorice 2013	Sorice P, Santoleri F, La Sala R, Rizzo C, Scurti V, Costantini A. Adherence and persistence in kidneys cancer oral therapy. *Int J Clin Pharm* 2013:968	Abstract with insufficient methodological details
Stillman 2013	Stillman I, Ambavane O, Cislo P. A cost-effectiveness analysis of axitinib and sorafenib for 2nd line treatment of advanced renal cell carcinoma after failure of cytokines in the United States. *Value Health* 2013:A410	Abstract with insufficient methodological details
Su 2012	Su Y, Shi N, Landsman-Blumberg P, Poehlein C, Waxman I. First-line targeted agents and cost of care for patients with metastatic renal cell cancer (MRCC). *Ann Oncol* 2012;**23**(Suppl.9):ix258–93	Abstract with insufficient methodological details
Sura 2012	Sura M, Goryaynov S, Avxentyeva M, Omelyanovsky V. Cost-mimimization analysis of pazopanib versus sunitinib, sorafenib and bevacizumab + interferon alpha-2 A for patients with metastatic renal cell carcinoma. *Value Health* 2012:A458	Abstract with insufficient methodological details
Swinburn 2010	Swinburn P, Lloyd A, Nathan P, Chouieri TK, Cella D, Neary MP. Elicitation of health state utilities in metastatic renal cell carcinoma. *Curr Med Res Opin* 2010;**26**:1091–6	Wrong population
Ta 2013	Ta AD, Bolton DM, Dimech MK, White V, Davis ID, Coory M, *et al.* Contemporary management of renal cell carcinoma (RCC) in Victoria: implications for longer term outcomes and costs. *BJU Int* 2013;**112**(Suppl. 2):36–43	No relevant outcomes
Tatar 2009	Tatar M, Akbulut H. Cost-effectiveness of sorafenib in unresectable and/or metastatic renal cell carcinoma in Turkey. *Value Health* 2009:A222	Abstract with insufficient methodological details
Tatokoro 2011	Tatokoro M, Fujii Y, Kawakami S, Saito K, Koga F, Matsuoka Y, *et al.* Phase-II trial of combination treatment of interferon-alfa, cimetidine, cyclooxygenase-2 inhibitor and renin-angiotensin-system inhibitor (I-CCA) for metastatic renal cell carcinoma. *Eur Urol* 2011:230	No relevant outcomes
Taylor 2010	Taylor P, Wing J, Mapp K, Pavlakis N, De Souza P, Grygiel K, Pezzullo L. The cost of side effects associated with treatment for advanced renal cell carcinoma in Australia. *Asia-Pacific J Clin Oncol* 2010;**6**:153	Abstract with insufficient methodological details
Teich 2009	Teich V, Fernandes RA, Schiola A. Cost-effectiveness analysis of sorafenib associated to best supportive care (BSC) versus best supportive care alone in the second line treatment of advanced renal cell carcinoma under the Brazilian public health care system perspective. *Value Health* 2009:A42	Abstract with insufficient methodological details
Teich 2010	Teich V, Hashizume CM, Marinho T, Charbonneau C, Naves A. Economic evaluation of sunitinib vs. interferon-a and bevacizumab + interferon-a in the treatment of metastatic renal cell carcinoma (CCRM)-brazilian private health system perspective. *Value Health* 201):A37	Abstract with insufficient methodological details
Tenorio 2009	Tenorio C, Vargas J, Rizo-Rios P, Flores-Gil O, Martínez-Fonseca J, Mould-Quevedo J, *et al.* Pharmacoeconomic evaluation of sunitinib malate for first-line treatment of metastatic renal cell carcinoma in Mexico. *Value Health* 2009:A44–5	Abstract with insufficient methodological details
Topibulpong 2010	Topibulpong N, Tanasanvimon S, Parinyanitikul N, Vinayanuwattikun C, Sriuranpong V. Economic implications of the first-line treatment of advanced renal cell carcinoma in Thailand: a cost-effectiveness analysis. *J Clin Oncol* 2010;**28**(Suppl.15):e15136	Abstract with insufficient methodological details

Study	Reference	Reason for exclusion
Topibulpong 2010	Topibulpong N, Tanasanvimon S, Parinyanitikul N, Vinayanuwattikun C, Sriuranpong V. Economic implications of the first-line treatment of advanced renal cell carcinoma in Thailand: a cost-effectiveness analysis. *J Clin Oncol* 2010	Abstract with insufficient methodological details
Torres Toala 2013	Torres Toala FG, Riofrio A, Mould JF, Estevez C. Cost-effectiveness and cost–utility analysis of sunitinib versus sorafenib and bevacizumab + interferon-alfa as first-line treatment for metastatic renal cell carcinoma in Ecuador. *Value Health* 2013;A139	Abstract with insufficient methodological details
Villa 2013	Villa G, Hernandez-Pastor L-J. Budget impact analysis of first-line treatment with pazopanib for advanced renal cell carcinoma in Spain. *BMC Cancer* 2013;**13**:399	Wrong population
Vogelzang 2013	Vogelzang NJ, Bhor M, Liu Z, Dhanda R, Hutson TE. Everolimus vs. temsirolimus for advanced renal cell carcinoma: use and use of resources in the US Oncology Network. *Clin Genitourin Cancer* 2013;**11**:115–20	Non-UK costing study
Wong 2013	Wong MK, Wang X, Chulikavit MJ, Liu Z. Review of US comparative economic evidence for treatment of metastatic renal cell carcinoma after failure of first-line VEGF inhibitor therapy. *American Health & Drug Benefit*. 2013;**6**:275–86	Systematic review
Wu 2012	Wu B, Dong B, Xu Y, Zhang Q, Shen J, Chen H, Xue W. Economic evaluation of first-line treatments for metastatic renal cell carcinoma: a cost-effectiveness analysis in a health resource-limited setting. *PLOS ONE* 2012;**7**:e32530	Wrong population
Wu 2011	Wu JZ. Economic evaluation of sunitinib malate for the first-line treatment of metastaric renal cell carcinoma in the Chinese health care setting. *Value Health* 2011;A454	Abstract with insufficient methodological details
Wu 2014	Wu XS. Budget impact model of sunitinib as first line treatment of metastatic renal cell carcinoma in China. *Value Health* 2014;A734	Abstract with insufficient methodological details
Yagudina 2012	Yagudina RK. Economic evaluation of sunitinib for the first-line treatment of metastatic renal cell carcinoma in Russian federation. *Value Health* 2012;A222	Abstract with insufficient methodological details
Zheng 2012	Zheng ZD, Qu SX, Liu YY, Hao H, Zhang GJ, Xie XD. Clinical controlled trial of first-line treatment for advanced kidney cancer. *Chung-Hua i Hsueh Tsa Chih* 2012;**92**:2984–7	Study not available in the English language

Excluded quality-of-life studies

Study	Reference	Reason for exclusion
Anonymous	Patient preference between pazopanib (Paz) and sunitinib (Sun): results of a randomised double-blind, placebo-controlled, cross-over study in patients with metastatic renal cell carcinoma (mRCC)-PISCES study, NCT 01064310. *Clin Adv Hematol Oncol* 2012;**10**:8–11	First line
Bushmakin 2012	Bushmakin A, Cappelleri JC, Korytowsky B, Sandin R, Matczak E, Cella D. Sunitinib (SU) dosing schedule and data collection timepoints: impact on quality of life (QOL) outcomes in metastatic renal cell carcinoma (MRCC). Conference: 37th ESMO Congress, Vienna, 2012. *Ann Oncol* 2012;**23**:ix269	First line
Cella 2009	Cella D, Cappelleri JC, Bushmakin A, Charbonneau C, Li JZ, Kim ST, *et al.* Quality of life predicts progression-free survival in patients with metastatic renal cell carcinoma treated with sunitinib versus interferon alfa. *J Oncol Pract* 2009;**5**:66–70	First line

Study	Reference	Reason for exclusion
Cella 2012	Cella D, Kaiser K, Beaumont J, Diaz J, McCann L, Mehmud F, *et al.* Quality of life (QOL) among renal cell carcinoma (RCC) patients in a randomised double blind cross-over patient preference study of pazopanib (P) versus sunitinib (S). Conference: 37th ESMO Congress, Vienna, 2012. *Ann Oncol* 2012;**23**:ix261–ix262	First line
De Groot 2014	De Groot S, Redekop W, Oosterwijk E, Kiemeney LC, Uyl-De Groot C. Patient and disease characteristics are important determinants of health-related quality of life of patients with metastatic renal cell carcinoma results from a population-based registry. Conference: ISPOR 17th Annual European Congress, Amsterdam, 2014. *Value Health* 2014;**17**:A649	Line of treatment not stated (probably a mix)
Elfiky 2009	Elfiky A, Oh WK, Choueiri TK. Health-related quality of life and symptom improvement for patients with metastatic renal cell carcinoma: evaluating sunitinib and interferon alfa. *Am J Hematol/Oncol* 2009;**8**	Irretrievable
Tannir 2006	Tannir NM, Cohen L, Wang X, Thall P, Mathew PF, Jonasch E, *et al.* Improved tolerability and quality of life with maintained efficacy using twice-daily low-dose interferon-alpha-2b: results of a randomised phase II trial of low-dose versus intermediate-dose interferon-alpha-2b in patients with metastatic renal cell carcinoma. *Cancer* 2006;**107**:2254–61	Non generic preference-based measure
Zbrozek 2010	Zbrozek AS, Hudes G, Levy D, Strahs A, Berkenblit A, DeMarinis R, *et al.* Q-TWiST analysis of patients receiving temsirolimus or interferon alpha for treatment of advanced renal cell carcinoma. *PharmacoEconomics* 2010;**28**:577–84	Interventions not of interest
Rixe 2007	Rixe O, Bukowski RM, Michaelson MD, Wilding G, Hudes GR, Bolte O, *et al.* Axitinib treatment in patients with cytokine-refractory metastatic renal-cell cancer: a phase II study. *Lancet Oncol* 2007;**8**:975–84	Disease-specific measure (EORTC QLQ-C30)
Motzer 2014	Motzer RJ, Escudier B, Tomczak P, Hutson TE, Michaelson MD, Negrier J, *et al.* Axitinib versus sorafenib as second-line treatment for advanced renal cell carcinoma: overall survival analysis and updated results from a randomised phase 3 trial. *Lancet Oncol* 2013;**14**:552–62	Non-generic preference-based measure (FKSI-DRS and FKSI-D15)
Kröger 1999	Kröger MJ, Menzel T, Gschwend JE, Bergmann L. Life quality of patients with metastatic renal cell carcinoma and chemo-immunotherapy – a pilot study. *Anticancer Res* 1999;**19**:1553–5	Non-generic preference-based measure

Appendix 11 WinBUGS code

WinBUGs code for fixed-effects model: hazard ratio

```
model{
for(i in 1:ndp){
 prec[i]<- 1/(se[i]*se[i])
lhr[i]~dnorm(md[i],prec[i])
md[i] <- d[t[i]] – d[b[i]]
dev[i] <- (lhr[i] – md[i])*(lhr[i] – md[i])/(se[i]*se[i])
}
resdev <- sum(dev[])
d[1]< –0
for (k in 2:nt){
d[k] ~ dnorm(0,.001)
}
for(k in 1:nt){
rk[k]<- rank(d[],k)
best[k]<-equals(rk[k],1)
}
for (c in 1:nt-1){
for (k in (c+ 1):nt){
lhzr[c,k] <- d[k] – d[c]
HR[c,k] <- exp(lhzr[c,k])
}
}
}
```

WinBUGS code for fixed-effects model: odds ratio

```
model{
for(i in 1:ns){
delta[i,t[i,1]]< –0
mu[i] ~ dnorm(0,.0001)
for (k in 1:na[i]) {
r[i,t[i,k]] ~ dbin(p[i,t[i,k]],n[i,t[i,k]])
logit(p[i,t[i,k]])<-mu[i] + delta[i,t[i,k]]
rhat[i,t[i,k]]<- p[i,t[i,k]] * n[i,t[i,k]]
resdev[i,k]<- 2 * (r[i,t[i,k]] * (log(r[i,t[i,k]]) – log(rhat[i,t[i,k]])) + (n[i,t[i,k]] – r[i,t[i,k]]) *
(log(n[i,t[i,k]] – r[i,t[i,k]]) – log(n[i,t[i,k]] – rhat[i,t[i,k]])))
}
sumdev[i]<-sum(resdev[i,1:na[i]])
for (k in 2:na[i]) {
delta[i,t[i,k]] <- d[t[i,k]] – d[t[i,1]]
}
}
sumdevtot<- sum(sumdev[])
d[1]< –0
for (k in 2:nt){
d[k] ~ dnorm(0,.0001)
}
for (i in 1:ns) {
mu1[i] <- mu[i] * equals(t[i,1],1)
}
```

```
for (k in 1:nt) {

logit(T[k])<- sum(mu1[])/nb +d[k]

}

for (k in 1:nt) {

rk[k]<-nt – rank(T[],k)

best[k]<-equals(rk[k],1)

}

for (c in 1:(nt-1)) {

for (k in (c+ 1):nt) {

lor[c,k] <- (d[k] – d[c])

or[c,k]<-exp(lor[c,k])

}

}

}
```

Appendix 12 Results of TA219

Results of the company's original analysis from TA219.

Treatment	Median PFS (months)		Mean OS (months)[a]		Total cost (£)	Total QALYs	ICER (£/QALY)
Everolimus + BSC	4.90 (95% CI 3.98 to 5.52)	HR: 0.33 (95% CI 0.25 to 0.43)	10.1	HR: 0.55 (95% CI 0.31 to 0.97)	NR	NR	51,613
BSC	1.87 (95% CI 1.84 to 1.94)		5.1		NR	NR	–

NR, not reported.
a Results of inverse proportion of censoring weights (IPCW) analysis.

Results of the final ERG base-case analysis from TA219.

Treatment	Median PFS (months)		Mean OS (months)	Total incremental cost (£)	Total incremental QALYs	ICER (£/QALY)
Everolimus + BSC	4.90 (95% CI 3.98 to 5.52)	HR: 0.33 (95% CI 0.25 to 0.43)	14.1	18,986	0.33	58,316
BSC	1.87 (95% CI 1.84 to 1.94)		8.9			–

Appendix 13 Mean EQ-5D results from Cella *et al.*[130]

Mean EQ-5D results from Cella *et al.*[130]

Time point	Treatment arm, score (number of patients)	
	Pazopanib	Placebo
Baseline	EQ-5D index: 0.72 ± 0.25 ($n = 287$)	EQ-5D index: 0.73 ± 0.24 ($n = 143$)
	EQ-5D VAS: 64.6 ± 23.69 ($n = 283$)	EQ-5D VAS: 65.9 ± 23.84 ($n = 141$)
Week 6	EQ-5D index: 0.71 ± 0.22 ($n = 255$)	EQ-5D index: 0.72 ± 0.30 ($n = 127$)
	EQ-5D VAS: 65.5 ± 21.84 ($n = 244$)	EQ-5D VAS: 64.7 ± 24.37 ($n = 115$)
Week 12	EQ-5D index: 0.70 ± 0.25 ($n = 221$)	EQ-5D index: 0.75 ± 0.23 ($n = 87$)
	EQ-5D VAS: 67.8 ± 20.89 ($n = 216$)	EQ-5D VAS: 68.6 ± 22.75 ($n = 82$)
Week 18	EQ-5D index: 0.71 ± 0.26 ($n = 197$)	EQ-5D index: 0.76 ± 0.22 ($n = 62$)
	EQ-5D VAS: 67.6 ± 20.18 ($n = 191$)	EQ-5D VAS: 68.4 ± 20.24 ($n = 61$)
Week 24	EQ-5D index: 0.71 ± 0.24 ($n = 168$)	EQ-5D index: 0.76 ± 0.23 ($n = 51$)
	EQ-5D VAS: 70.8 ± 17.32 ($n = 164$)	EQ-5D VAS: 70.4 ± 19.5 ($n = 49$)
Week 48	EQ-5D index: 0.79 ± 0.20 ($n = 98$)	EQ-5D index: 0.80 ± 0.24 ($n = 24$)
	EQ-5D VAS: 72.0 ± 17.78 ($n = 95$)	EQ-5D VAS: 73.1 ± 17.29 ($n = 23$)

Appendix 14 Prevalence of adverse events

Summary of treatment-related adverse events/treatment-emergent adverse events used in the model

Treatment (source of estimate)	TRAEs ≥ grade 3
Axitinib (TA333)[89]	Hypertension: 15.3%
	Diarrhoea:10%
Nivolumab (TA417)[90]	Pneumonitis: 1.5%
	Diarrhoea: 1.2%
	Anaemia: 1.7%
Everolimus (TA417)[42]	Pneumonitis: 2.8%
	Diarrhoea: 1.3%
	Anaemia: 7.8%
Everolimus (TA219)[42]	Anaemia: 10.2%
	Anorexia/cachexia: 1.5%
	Nausea/vomiting: 3.7%
	Dyspnoea: 7.7%
	Infections: 3%
	Pneumonitis single term: 2.6%
[a]Cabozantinib (TA463)[28]	Diarrhoea: 13.0%
	Anaemia: 5.7%
	Hypertension: 14.8%
	PPE: 4.0%

PPE, palmar–plantar erythrodysesthesia.
a Only data on TEAEs were identified for cabozantinib.

Appendix 15 Goodness-of-fit statistics

Goodness-of-fit statistics for nivolumab fitted overall survival curves

Distribution	AIC	BIC
Exponential	1749.19	1753.21
Weibull	1737.94	1745.98
Log-normal	1745.25	1753.28
Log-logistic	1738.66	1746.69
Gompertz	1742.17	1750.20
Generalised gamma	1739.71	1751.76
Generalised *F*	N/A[a]	1757.80
Spline (one knot)	1739.78	1751.83
Spline (two knot)	1741.74	1757.80

a The generalised *F* model had not converged so an AIC value was not generated.

Goodness-of-fit statistics for everolimus fitted overall survival curves

Distribution	AIC	BIC
Exponential	1908.88	1912.90
Weibull	1907.63	1915.67
Log-normal	1902.59	1910.62
Log-logistic	1905.54	1913.58
Gompertz	1910.10	1918.14
Generalised gamma	1904.04	1916.10
Generalised *F*	1906.05	1922.12
Spline (one knot)	1904.77	1916.83
Spline (two knot)	1901.86	1917.93

Goodness-of-fit statistics for nivolumab fitted progression-free survival curves

Distribution	AIC	BIC
Exponential	2051.71	2055.72
Weibull	2051.09	2060.12
Log-normal	1960.78	1968.81
Log-logistic	1976.43	1984.46
Gompertz	2020.65	2028.68
Generalised gamma	1878.81	1890.86
Generalised *F*	N/A[a]	N/A[a]
Spline (one knot)	1930.61	1942.66
Spline (two knot)	1864.21	1880.27

a The generalised *F* model had not converged so an AIC and BIC values were not generated.

Goodness-of-fit statistics for everolimus fitted progression-free survival curves

Distribution	AIC	BIC
Exponential	1939.69	1943.71
Weibull	1937.73	1945.77
Log-normal	1880.58	1888.61
Log-logistic	1895.82	1903.86
Gompertz	1940.31	1984.35
Generalised gamma	1875.68	1887.73
Generalised *F*	1877.69	1893.77
Spline (one knot)	1884.64	1896.70
Spline (two knot)	1855.17	1871.24

Goodness-of-fit statistics for everolimus fitted time to discontinuation curves

Distribution	AIC	BIC
Exponential	2383.30	2387.32
Weibull	2384.10	2392.13
Log-normal	2316.16	2324.20
Log-logistic	2334.04	2342.08
Gompertz	2381.77	2389.80
Generalised gamma	2309.99	2322.05
Generalised *F*	2312.00	2328.08
Spline (one knot)	2319.02	2331.07
Spline (two knot)	2292.91	2308.98

Goodness-of-fit statistics for nivolumab fitted time to discontinuation curves

Distribution	AIC	BIC
Exponential	2568.58	2572.60
Weibull	2570.34	2578.37
Log-normal	2524.15	2532.19
Log-logistic	2534.37	2542.40
Gompertz	2565.86	2573.90
Generalised gamma	2525.10	2537.15
Generalised *F*	2527.11	2543.17
Spline (one knot)	2524.16	2536.20
Spline (two knot)	2520.81	2536.87

Goodness-of-fit statistics for cabozantinib fitted time to discontinuation curves

Distribution	AIC	BIC
Exponential	2026.29	2030.09
Weibull	2024.59	2032.20
Log-normal	1987.44	1995.04
Log-logistic	1989.29	1996.89
Gompertz	2025.96	2033.56
Generalised gamma	1987.60	1999.01
Generalised *F*	1987.64	2002.85
Spline (one knot)	1991.09	2002.49
Spline (two knot)	1990.84	2006.05

Appendix 16 Probabilistic analysis parameters

The PSA parameter distributions probabilistic sensitivity analyses

Category	Parameter	Distribution	Value	Alpha	Beta
Treatment costs	Nivolumab administration	Gamma	£186	7.012	26.460
Disease costs	Consultant visit	Gamma	£189	5.720	32.978
Disease costs	CT scan	Gamma	£136	3.506	38.850
Disease costs	Blood test	Gamma	£3	4.341	0.692
Disease costs	Specialist community nurse visit	Gamma	£75	16	4.688
Disease costs	Pain medication	Gamma	£5	16	0.328
TC costs	District nurse	Gamma	£67	16	4.188
TC costs	Hospital costs	Gamma	£6239	16	389.938
TC costs	LA care costs	Gamma	£470	16	29.375
Subsequent treatment cost	Cabozantinib ST	Gamma	£4476	16	279.764
Subsequent treatment cost	Axitinib ST	Gamma	£5878	16	367.364
Subsequent treatment cost	Everolimus ST	Gamma	£3301	16	206.306
Subsequent treatment cost	Nivolumab ST	Gamma	£5494	16	343.348
Resource use disease	Consultant visits per week	Gamma	0.08	16	0.005
Resource use disease	CT scans per week	Gamma	0.08	16	0.005
Resource use disease	Blood tests per week	Gamma	0.25	16	0.016
Resource use disease	Specialist community nurse visits per week	Gamma	0.38	16	0.024
Resource use disease	Pain medications per week	Gamma	7	16	0.438
Resource use TC	GP visits per cycle	Gamma	11.4	16	0.713
Resource use TC	District nurse visits per cycle	Gamma	7.5	16	0.469
Resource use TC	Hospital care per cycle	Gamma	1	16	0.063
Resource use TC	LA care per cycle	Gamma	1	16	0.063
Resource use AE	GP visits per cycle for hypertension	Gamma	0.08	16	0.005
Resource use AE	District nurse visits per cycle for hypertension	Gamma	0.08	16	0.005
Resource use AE	Medication for hypertension per cycle	Gamma	14	16	0.875
Resource use AE	GP visit for pneumonitis per cycle	Gamma	2	16	0.125
Resource use AE	Steroids for pneumonitis	Gamma	1	16	0.063
Resource use AE	Consultant appointments for diarrhoea (nivolumab)	Gamma	1	16	0.063
Resource use AE	Loperamide for diarrhoea (nivolumab)	Gamma	1	16	0.063
Resource use AE	Regular day and night admission anaemia	Gamma	1	16	0.063
Resource use AE	Radiotherapy BSC	Gamma	1	16	0.063

Category	Parameter	Distribution	Value	Alpha	Beta
Resource use AE	PPE	Gamma	1	16	0.063
AE cost	Regular day and night admission anaemia	Gamma	£422	2.238	188.356
AE cost	Radiotherapy BSC treatment	Gamma	£109	11.521	9.452
AE cost	Radiotherapy BSC planning	Gamma	£310	9.172	33.763
AE cost	PPE	Gamma	£101	16	6.313
Dose adjustments	Axitinib	Gamma	100%	16	0.064
Dose adjustments	Nivolumab	Gamma	100%	16	0.063
Dose adjustments	Everolimus	Gamma	100%	16	0.063
Dose adjustments	Cabozantinib	Gamma	100%	16	0.063
Dose adjustments	Sunitinib	Gamma	100%	16	0.063
AE probability	Hypertension (axitinib)	Beta	0.153	55	304
AE probability	Diarrhoea (axitinib)	Beta	0.100	36	323
AE probability	Diarrhoea (cabozantinib)	Beta	0.050	17	314
AE probability	Anaemia (cabozantinib)	Beta	0.040	13	318
AE probability	Hypertension (cabozantinib)	Beta	0.080	26	305
AE probability	PPE (cabozantinib)	Beta	0.040	13	318
AE probability	Pneumonitis (nivolumab)	Beta	0.015	6	400
AE probability	Diarrhoea (nivolumab)	Beta	0.012	5	401
AE probability	Anaemia (nivolumab)	Beta	0.017	7	400
AE probability	Pneumonitis (everolimus)	Beta	0.028	11	386
AE probability	Diarrhoea (everolimus)	Beta	0.013	5	392
AE probability	Anaemia (everolimus)	Beta	0.078	31	366
AE probability	Radiotherapy (BSC)	Beta	0.100	10	90
Utilities	PFS	Beta	0.692	500.316	222.684
Utilities	PPS	Beta	0.610	441.030	281.970
Utilities	PFS (nivolumab)	Beta	0.728	526.344	196.656
Utilities	PPS (nivolumab)	Beta	0.646	467.058	255.942

LA, local authority; PPE, palmar–plantar erythrodysesthesia; ST, subsequent treatment.

Appendix 17 Undiscounted cost-effectiveness results

Summary of costs (undiscounted)

	Treatment				
Cost (£) component	Cabozantinib	Axitinib	Everolimus	Nivolumab	BSC
PFS costs					
Treatment acquisition	80,325	29,930	20,675	79,968	0
AE	724	23	589	250	250
Disease management (on treatment)	1864	948	923	1281	332
Disease management (off treatment)	53	0	25	23	0
Total PFS	82,966	30,901	22,212	81,523	582
PPS costs					
Treatment acquisition	14,984	0	337	13,213	0
AEs	135	0	10	41	0
Disease management (on treatment)	819	0	35	499	0
Disease management (off treatment)	5496	5292	5256	5047	3482
Subsequent therapy	2476	4350	2450	3624	0
End of life	7713	7713	7713	7713	7713
Total PPS	31,623	17,335	15,801	30,137	11,195
Total costs	114,589	48,256	38,013	111,660	11,777

Summary of quality-adjusted life-years (undiscounted)

	Treatment				
Health state	Cabozantinib	Axitinib	Everolimus	Nivolumab	BSC
PFS QALYs					
On treatment	0.89	0.45	0.44	0.64	0.16
Off treatment	0.03	0.00	0.01	0.01	0.00
Total PFS	0.92	0.45	0.45	0.66	0.16
PPS QALYs					
On treatment	0.15	0.00	0.01	0.09	0.00
Off treatment	0.98	0.95	0.94	0.96	0.62
Total PPS	1.13	0.95	0.95	1.05	0.62
Total QALYs	2.05	1.40	1.40	1.71	0.78

Incremental cost-effectiveness results (undiscounted)

Incremental analysis	Total cost (£)	Total QALYs	Total life-years	ICER (£)
BSC	11,777	0.78	1.25	–
Everolimus	38,013	1.40	2.21	42,456
Axitinib	48,256	1.40	2.21	Dominated by everolimus
Nivolumab	111,660	1.71	2.53	Dominated (extended) by cabozantinib
Cabozantinib	114,589	2.05	3.18	118,521

Appendix 18 Results of scenario analyses

Scenario 1: using alternative hazard ratios

Incremental analysis	Total cost (£)	Total QALYs	Total life-years	ICER (£)
BSC	11,306	0.75	1.25	–
Everolimus	36,463	1.31	2.21	44,987
Axitinib	46,151	1.03	1.68	Dominated by everolimus
Cabozantinib	106,559	1.88	3.19	124,471
Nivolumab	107,838	1.80	2.92	Dominated by cabozantinib

Scenario 2: using the Gompertz distribution for overall survival

Incremental analysis	Total cost (£)	Total QALYs	Total life-years	ICER (£)
BSC	11,255	0.74	1.23	–
Everolimus	36,269	1.28	2.14	46,481
Axitinib	46,289	1.28	2.14	Dominated by everolimus
Nivolumab	104,961	1.43	2.23	Dominated (extended) by cabozantinib
Cabozantinib	105,170	1.78	2.99	137,487

Scenario 3: using the exponential distribution for overall survival

Incremental analysis	Total cost (£)	Total QALYs	Total life-years	ICER (£)
BSC	11,357	0.76	1.27	–
Everolimus	36,907	1.41	2.40	39,650
Axitinib	46,909	1.41	2.40	Dominated by everolimus
Cabozantinib	107,642	2.07	3.61	107,615
Nivolumab	108,257	1.90	3.14	Dominated by cabozantinib

Scenario 4: using the one-knot spline for overall survival

Incremental analysis	Total cost (£)	Total QALYs	Total life-years	ICER (£)
BSC	11,430	0.78	1.30	–
Everolimus	37,163	1.45	2.48	38,487
Axitinib	47,203	1.45	2.48	Dominated by everolimus
Nivolumab	106,907	1.62	2.58	Dominated (extended) by cabozantinib
Cabozantinib	108,133	2.14	3.79	101,889

Scenario 5: using the two-knot spline for overall survival

Incremental analysis	Total cost (£)	Total QALYs	Total life-years	ICER (£)
BSC	11,241	0.74	1.23	–
Everolimus	36,314	1.29	2.15	46,011
Axitinib	46,347	1.29	2.15	Dominated by everolimus
Cabozantinib	106,094	1.82	3.07	131,643
Nivolumab	107,018	1.64	2.61	Dominated by cabozantinib

Scenario 6: using the log-normal distribution for time to discontinuation

Incremental analysis	Total cost (£)	Total QALYs	Total life-years	ICER (£)
BSC	11,304	0.75	1.25	–
Everolimus	36,586	1.31	2.21	45,183
Axitinib	46,506	1.31	2.21	Dominated by everolimus
Cabozantinib	103,469	1.87	3.18	120,519
Nivolumab	109,797	1.60	2.53	Dominated by cabozantinib

Scenario 7: assuming overall survival for cabozantinib is equivalent to nivolumab

Incremental analysis	Total cost (£)	Total QALYs	Total life-years	ICER (£)
BSC	11,304	0.75	1.25	–
Everolimus	36,463	1.31	2.21	44,965
Axitinib	46,506	1.31	2.21	Dominated by everolimus
Cabozantinib	102,922	1.54	2.53	Dominated (extended) by nivolumab
Nivolumab	106,761	1.60	2.53	247,971

Scenario 8: treatment discontinuation at the point of progression

Incremental analysis	Total cost (£)	Total QALYs	Total life-years	ICER (£)
BSC	11,304	0.75	1.25	–
Everolimus	36,701	1.31	2.21	45,390
Axitinib	46,506	1.31	2.21	Dominated by everolimus
Nivolumab	95,290	1.60	2.53	Dominated (extended) by cabozantinib
Cabozantinib	96,555	1.87	3.18	107,852

Appendix 19 One-way sensitivity analysis results

Key one-way sensitivity analysis results (change from base case)

Parameters	Treatment									
	Everolimus		Nivolumab		Axitinib		Cabozantinib		BSC	
	ΔCost (£)	ΔQALYs	ΔCost (£)	ΔQALYs	ΔCost (£)	ΔQALYs	ΔCost (£)	ΔQALYs	ΔCost (£)	ΔQALYs
OS HR (lower)	–	–	–	–	–	–	2265	0.36	5936	1.15
OS HR (upper)	–	–	–	–	–	–	–2239	–0.28	–1991	–0.39
Relative dose adjustments (lower)	–8791	–	–36,291	–	–12,203	–	–38,118	–	–	–
Relative dose adjustments (upper)	11,209	–	46,274	–	16,891	–	48,604	–	–	–
Nivolumab administration costs (lower)	–	–	–2854	–	–	–	–	–	–	–
Nivolumab administration costs (upper)	–	–	4130	–	–	–	–	–	–	–
TC hospital costs (lower)	–2483	–	–2456	–	–2483	–	–2407	–	–2562	–
TC hospital costs (upper)	3166	–	3131	–	3166	–	3069	–	3267	–
Subsequent treatment costs (lower)	–1029	–	–1516	–	–1825	–	–1031	–	–	–
Subsequent treatment costs (upper)	1312	–	1933	–	2328	–	1314	–	–	–
PFS utilities (lower)	–	–0.02	–	–0.03	–	–0.02	–	–0.04	–	–0.01
PFS utilities (upper)	–	0.02	–	0.03	–	0.02	–	0.04	–	0.01
PPS utilities (lower)	–	–0.05	–	–0.05	–	–0.05	–	–0.06	–	–0.03
PPS utilities (upper)	–	0.05	–	0.05	–	0.05	–	0.06	–	0.03

The net monetary benefit for key one-way sensitivity analysis

	Treatment (£)				
NMB	Everolimus	Nivolumab	Axitinib	Cabozantinib	BSC
£20,000 per QALY threshold					
Base case	–10,200	–74,828	–20,243	–69,153	3769
No OWSAs changed the ranking of NMB					
£30,000 per QALY threshold					
Base case	*2931*	*–58,861*	*–7112*	*–50,472*	*11,305*
OS HR (upper)	2931	–58,861	–7112	–56,751	1705
Everolimus RDI (lower)	11,722	–58,861	–7112	–50,472	11,305
Everolimus RDI (upper)	–8277	–58,861	–7112	–50,472	11,305
Axitinib RDI (lower)	2931	–58,861	5091	–50,472	11,305
Nivolumab RDI (lower)	2931	–22,570	–7112	–50,472	11,305
Cabozantinib RDI (upper)	2931	–58,861	–7112	–99,077	11,305

EME
HS&DR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

Published by the NIHR Journals Library